Self assembly of collagen mimetic peptides

A thesis submitted in partial fulfillment of the requirements

for the

Degree of Master of Science in Biochemistry

2017

Deepti Mahapatra

University of Canterbury

Acknowledgements

I would like to start by thanking, all my four supervisors. This thesis would not be possible without Juliet, Ren, Jolon and Celine. Apart from being a motherly supervisor, Juliet stood beside me, and my work throughout. Her multitasking and work ethics are huge inspiration to me. I will never forget the importance of time management I learnt from Ren and thanks to his consistent pushes, I have evolved at my work today. Jolon's optimism and disciplines throughout my research work have been infections. Celine's presence and ability to simply and explain complicated concepts have taught me a lot from day 1. Thanks BIC and AgResearch for funding my project and my scholarship. Thanks to my supervisors for funding many domestic, international conferences, synchrotron trips and workshops.

Thank Jackie and Rayleen for training me at many instruments in the lab. Thanks Jackie for her help with TEM images and thanks Luigi for his guidance with AFM. Thanks Helen and Gary from Electrical Engineering for teaching me AFM. Thanks Peter for his kind support with my racemic crystallography work. Thanks Justine Cottam and Matthew Polson for their time. Thanks Gert Jan, Logan and the people from chemistry for helping out at various instances. Thanks Dr Ali Nazmi and Dimitri for making me familiar with mosquito and taking crystal pictures, initially.

Thanks people at Australian Synchrotron (Melbourne), for their valuable assistance at IR, SAXS, MX1, MX2 beamlines; Especially Mark and Lili at the IR beam line.

Thanks AgResearch for being a generous collaborator. I am glad to have learnt Redox proteomics from Dr. Santau Deb-Choudhury and my interactions with Anita, James, Stefan, Robert, Jou and everyone from Jolon's team have been an amazing experience.

Thanks Dr. Hironori Suzuki for teaching me crystallography. Thanks Dr. Celine Valery for teaching me self-assembly basics. Thanks to Moritz, Susie, Amy Phillips and Madhu and for going extra miles & helping me settle with my initial PhD and proposal framing days. I have learnt a lot from, my consistently interactions with Celine, Hiro, Ali, Muge, Luigi, Sarah, Rishi, Arvind, Akshita, Kat, Rachel, Jen, Nivas, Manmeet, Amy Yewdall, and everybody from Gerrard's lab, Dobson's lab and the whole 6th floor, School of Biological Sciences. University of Canterbury has been wonderful and treated me like family at New Zealand, right from day 1.

Thanks Sitara, Sonia, Ayelen and my flatmates at Kahu, for being great friends through ups and downs of the 4 years of my PhD. Thanks to my brother, sister in law, Nishtha, Mausiji and Mamaji for their undying support throughout. Big thanks to my mother, father and Gurudev for their unconditional love, care and guidance.

(I dedicate my thesis to my nanaji, who passed away in July 2014)

Contents

Abstract

This project is part of a wider research program that contributes to and focuses on the utilization of lower value red meat components to create new meat-derived foods and ingredients. There have been significant advances in studying primary, secondary, tertiary and quaternary structures of food proteins in recent years but still, there are fundamental gaps in understanding the interactions of proteins and peptides in whole and reconstituted foods.

Collagen has a profound impact on food texture and this thesis focuses on analyzing collagen at the molecular level using collagen model peptides, to better understand self-assembly in meat derived food products. Short peptides are easy to design and synthesize, making them an excellent model system for studying self-assembly. Extensive literature survey has been done to understand self-assembly with short peptides, in Chapter 1. The materials and the methodologies adopted in this thesis have been dealt with in Chapter 2.

In Chapter 3 the effect of varying chain lengths of GPO peptides on triple helical formations and triple helical stabilities have been studied. After the characterization of triple helical assemblies and supramolecular assemblies, the most significant findings were - (1) the triple helix formation was found to be directly proportional to chain length of the peptide, concentration of mixture and time of incubation (2) longer chain lengths of peptides exhibited higher triple helical stabilities (3) triple helical formations and stabilities were supported in high salt environments (4) pH 7 favoured triple helical stability the most followed by pH 3.5 (5) self-association was observed to be limited by the low concentration of monomers.

The effect of temperature cycles on self-assembly of collagen peptides has been studied in Chapter 4. Infrared (IR) spectroscopy was used to study the changes brought about in the underlying polyproline II chains and in intra and inter hydrogen bonding patterns of the supramolecular assemblies and crystals of the GPO peptides. From techniques like AFM & DLS, heat treatment was seen to cause reduction in size of the GPO peptide aggregations. It is speculated that the shrinkage of these supramolecular assemblies were a result of breakage of the intra and inter helical H bonds after heat treatment. D-spacing in $(GPO)_{10}$ and $(GPO)_7$ peptides were found to be exactly the same, 12.26 Å, which indicates that there were similarities in packing for both triple helices. The D-spacing in the $(GOO)_{10}$ peptides was ~ 12.36 Å that suggests a different staggered pattern for the helical arrangement and hence a different supramolecular assembly pattern.

Examination of triple helical packing of GPO peptides with crystallography has been conducted in Chapter 5. Collagen model peptides of varying GPO chain lengths have been successfully crystallized. Racemic $(GPO)_5$-Gly peptides have been studied. The X-ray crystal structure of the $(GPO)_{10}$ peptide is solved at a resolution of 0.89 Å. Apart from generic features seen in the previously solved structures, detailed findings from this structure were - (1) proline rings can adopt two stable conformational states (2) the helical pitch was seen to vary around a 7/2 helix and (3) the stabilization of the crystalline network was water-mediated via hydroxyproline hydration, which bridges hydrogen bonding of helices.

The effect of oxidation on triple helices and supramolecular assemblies has been studied in chapter 6. $(GOO)_{10}$ peptides were characterized to understand the differences occurring in the assembly properties of GPO and GOO collagen peptides.

This was done to help explore possibilities of improving the food matrices. $(GOO)_{10}$ peptides were found to be more stable than the $(GPO)_{10}$ peptides. In second part of the chapter, the impact of heating (cooking) on self assembly of $(GPO)_7$ collagen model peptides has been investigated. Tandem mass spectrometric analysis has been carried out to sequence these peptides and characterize the oxidative modifications occurring in amino acid sites in the peptide chain. Middle parts of the $(GPO)_7$ peptide sequence have been found to be highly prone to oxidation.

Apart from bridging gaps in studies of oxidative modifications in red meat processing, in the food industry, this research on self assembly of collagen model peptides will also prove beneficial in medical applications like drug delivery, tissue engineering, biomedical applications related to collagen etc.

Abbreviations

(ΔG)	gibbs free energy
%	percentage
°C	degree Celsius
1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
Å	angstrom
Abs	absorbance
AFM	atomic force microscopy
Aw	error factor
BSA	bovine serum albumin
CD	circular dichroism
C_GO_G	C=O in glycine
C_XO_X	C=O in amino acid X
C_YO_Y	C=O in amino acid Y
d	diameter
dH	enthalpy
DHBC	di-hydroxy benzoic acid
DLS	dynamic light scattering
DSC	differential scanning calorimetry
DTT	dithiothreitol
FTIR	fourier transform infrared spectroscopy
GOO	glycine-4Rhydroxyproline -4Rhydroxyproline
GPO	glycine-proline-4Rhydroxyproline
(GPO)n	glycine-proline-4Rhydroxyproline; n is the number of repeats
GPP	glycine-proline-proline
HCCA	α-cyano-4-hydroxy-cinnamic acid
HPLC	high performance liquid chromatography
IAM	iodoacetamide
IR	infrared
IRM	infrared micro spectroscopy

KJ	kilo joule
L	path length of the cell (in cm)
MP	melting point
MRE	mean residual ellipticity
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MW	molecular weight
NMR	nuclear magnetic resonance
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffer saline
PDB	protein data bank
PEG	polyethylene glycol
pH	negative decadic logarithm of the hydrogen ion concentration
pKa	negative decadic logarithm of the acid dissociation constant
POG	proline- glycine-4Rhydroxyproline
RMSD	root mean square deviation
SAXS	small angle X-ray scattering
SDS	sodium dodecyl sulfate
TCEP	tris-2-carboxyethyl phosphine 2 propanol
TEM	transmission electron microscopy
T_m	transition temperature
TOF	time of flight
UV	ultra violet
ΔC_p	heat capacity change
ΔH	change in enthalpy
θ (λ)	molar ellipticity deg cm^2 $decimol^{-1}$
θ obs	observed ellipticity in degrees

1 Self assembly of collagen mimetic peptides

1.1 Introduction

The understanding of structural and biological functions of collagen can be enhanced by investigating self-assemblies of collagen model peptides [1]. This would further better the understanding of self-assembly in meat derived food products.

This research is focused on analyzing collagen at the molecular level using collagen mimetic peptides, to understand self-assembly and how the modification of collagen mimetic peptides might be used to control self-assembly.

This project is part of a wider research programme that focuses on utilization of lower value red meat components to create new meat derived foods and ingredients.

1.1.1 Oxidative modifications and self-assembly

Texture and tenderness are two key quality attributes influencing meat eating quality for consumers. Protein composition, modification and bimolecular architecture play a n important role in determining these qualities [2].

The change in properties associated with self-assembly owing to heat treatments and oxidative modifications have been studied in this research work. It has been hypothesized that differing textures can also be triggered by oxidative modifications Redox proteomic profiling approaches have also been utilized. The heat treatment steps have been performed to mimic the most common method of food processing, hydrothermal processing (cooking of meat), which brings changes in the texture of meat or meat derived food products.

1.1.2 Collagen and collagen mimetic peptides

Naturally abundant, the most common source of clinical collagen is of bovine, *Bos Taurus* [3]. Collagen, in its native form has not been used in this research because of issues like collagen can bring immunological and pathological effects with it if transplanted into humans [4, 5]. Collagen's intrinsic plasticity allows it to be developed as a biomaterial [3]. Owing to huge number of amino acid residues in native collagen, not many self-assembly experiments have been successful. Synthetic chemistry of short collagen model peptides, however, has proven to be more useful to study native collagen's self-assembly [6].

Collagen is an important muscle protein class and a major determinant of the physical properties of meat. It is also the most abundant connective tissue protein and is a contributing factor to the variation in meat tenderness and texture [7]. It is to be kept in mind that meat tenderness acts as the major factor affecting the consumer's assessment of meat quality [2]. The above factors have been encouraging to study collagen as a precursor to enhance the understanding of the quality of meat. Collagen model peptides have been a representative subject. In this research work, protein chemistry techniques have been used, to characterize collagen model peptides that form triple helical first order assemblies and self-associated higher order structures (supramolecular assemblies).

The current chapter discusses the factors that impact on the quality of red meat, mostly collagen. The following section outlines the important factors that determine the quality of meat.

1.2 Meat quality

1.2.1 Major components of red meat:

Commonly found proteins have been studied in detail e.g. actin [8], myosin [8], elastin [9] and collagen [2]. Apart from these proteins, there are various other meat-based bioactive substances that have been also been studied previously e.g. taurine (110 mg/100 g in lamb and 77 mg/100 g in beef) [10] and carnitine (209 mg/100 g in sheep [11] and 60 mg/ g in beef [12]), ubiquinone, glutathione, lipoic acid, spermine, carnosine, anserine [13].

1.2.2 Parameters influencing meat quality

The quality of meat is determined by its water holding capacity, texture, flavor, and its nutritional value [14-16]. According to consumer satisfaction surveys, meat quality is determined by the flavour of meat, as well as its juiciness and tenderness [2].

1.2.2.1 Texture

Tenderness is known to be determined by many factors including the age of the animal [17], diet, intramuscular fat, meat moisture content and animal slaughter conditions. The chemical and physical changes taking place in the muscle as it is processed, controls the level of tenderness in the final product. The important changes that take place arise due to the intermolecular and intramolecular cross-linking reactions between amino acids in the meat

proteins [18]. Intramolecular bonds are formed between amino acid residues within a single polypeptide chain or between two or more polypeptide chains within a protein, whereas intermolecular cross links are the bonds formed between two protein molecules [19]. The cross-links of collagen are of particular interest in this research, and have been discussed in detail in section 1.3.

1.2.2.2 Water holding capacity

The ability of meat to retain water is known as water holding capacity [20]. Amino acid composition and the amino acid sequence play an important role in meat quality, but apart from the distribution of amino acids, the aqueous environment plays an important role in muscle foods [14, 21]. Post slaughter, most of the water is held between thick and thin filaments of muscle cells and some of it is held between muscle proteins due to electrostatic attraction [20, 22]. During meat processing e.g. cooking, water loss can be a common problem. Cooked meat contains around 28-36% of proteins, whereas raw meat contains around 20-25% (Williams 2007). Upon cooking, the water content of meat decreases thereby increasing the protein content. It has been noted that loss of water holding capacity in oxidized proteins, such as collagen, due to cooking may lead to decreased juiciness in meat [15].

1.2.3 Role of collagen in meat quality

Meat is composed of various kinds of proteins including: myofibrillar proteins stromal proteins and sarcoplasmic proteins. Myofibrillar proteins include myosin, actin, troponin, tropomyosin, desmin, synemin, actinin, nebulin and numerous structural proteins. Stromal proteins are present in the connective tissue. Collagen and elastin are the major connective tissue proteins. Sarcoplasmic proteins are water-soluble. They form the intracellular fluid, and include enzymes that are important for growth and energy [23].

1.3 Insight into function and structure of collagen

1.3.1 Function of Collagen

Collagen is important in maintaining the structural integrity of tissues and organs. It acts as a backbone for "connective tissue", bone and cartilage in particular [24, 25]. It also regulates the biomechanical properties required for the proper functioning of these systems [24, 26]. Collagen is involved in the storage and release of cellular mediators, such as growth factors [24]. Any repair process of connective tissue requires composition and supramolecular

organization of a collagen matrix [18], e.g. fracture healing or treatment of bone, non-unions after trauma-surgery, or of cartilage defects etc [24]. It also possesses interesting features of biodegradability, immunogenicity, large-scale isolation possibilities and therefore has widespread usefulness in the medical, cosmetic and food industries [24].

The widespread distribution of collagen in the extracellular matrix of muscle gives structural support to the cells [27]. It is known that cooking of meat might bring changes in these widespread collagen networks, thus bringing changes in the post slaughter meat quality. This research, therefore, seeks to understand the relationship between collagen assembly and the properties of the assembled networks, in order to probe the effect on the quality of finished meat products.

1.3.2 Collagen structural characteristics

The types, structure and nature of the stabilizing cross-links in collagen play an important role in regulating the meat properties [2, 21].

1.3.2.1 Types of collagen

29 types of genetically distinct collagen have been described in literature [24]. Based on the structure and supramolecular organisation, collagen can be classified into five groups [28]:

(1) Fibril forming
(2) Fibril associated containing interrupted triple helices (FACIT)
(3) Beaded filament
(4) Anchoring fibril
(5) Network forming and trans-membrane filaments.

This thesis work has been focused on studying the textural changes of meat-derived foods via assemblies of collagen mimetic peptides. Focus being more on the fibril forming collagen. This collagen class comprises of collagen types I, II, III, V, and XI [24]. There are three major tracts of muscle connective tissue - epimysium, perimysium and endomysium. Both epimysium and perimysium comprise type I collagen as the most abundant protein and endomysium consists of collagens types I and III [29]. Type I collagen in muscular tissue has been the focus, due to its abundance in animal meat muscle.

1.3.2.2 Structure of collagen:

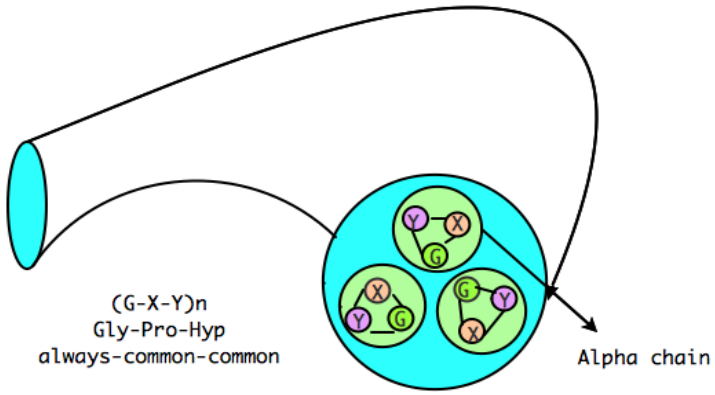


Figure 1.1 Structure of collagen triple helix, Figure adapted from [30] and drawn in ConceptDraw PRO

After translation, α-peptide chain that has a signal sequence at the N-terminus undergoes modifications in the endoplasmic reticulum. This peptide sequence is called the α-peptide or the α-chain [24].

This α-peptide then undergoes three types of modifications [24, 31]:

(1) Deletion of the signal sequence from the N-terminus;

(2) Hydroxylation of proline and lysine residues, forming hydroxyproline and hydroxylysine respectively, which aids in cross-linking in a later stage;

(3) Glycosylation (addition of glucose or galactose monomers) of the hydroxylysines but not the hydroxyproline residues.

The presence of the small amino acid glycine in every third residue along the chain facilitates winding [31] in the collagen triple helix. These three chains have a repeating structure of Gly-Xaa-Yaa, in which, Xaa and Yaa can be any amino acids but mostly are the amino acids, proline and hydroxyproline [31], as shown in Figure 1.1. There are three amino acids generally arranged in a triad, POG (P-proline, O-hydroxyproline, G-glycine).

When the three α-chains forming a triple helix are similar, it is referred to as a homotrimer, whereas three different α-chains form a heterotrimer. Typically, two similar α-chains and a different α-chain form a heterotrimer [32]. Homotrimers and heterotrimers play an important role in collagen structure and have been dealt with, in detail in sections 3 and 4.

After the above three post-translational modifications are complete, the α-strand is known as the propeptide. The propeptides bind together and the resulting triple helix is called procollagen [33] (Figure 1.2).

Procollagen consists of a 3000 Å triple helical domain. It contains three α-chains twisted in the form of helix and two globular trimeric propeptide domains on each side of the helix, at the N-terminus and the C-terminus. Once procollagen is secreted from the cell, it is converted into collagen by the removal of N and C propeptide domains by procollagen N-proteinase and procollagen C-proteinase respectively [34]. Collagen then self-assembles into cross-striated fibrils in extracellular connective tissue.
Stabilization of the collagen structure is facilitated by covalent cross-linking, which occurs by the oxidative deamination of specific lysine and hydroxylysine residues by lysyl oxidase [35-38].

1.3.2.3 Cross-links
Collagen cross-linking involves two basic pathways, namely: (i) the lysine aldehyde, or allysine cross-linking pathway and (ii) the hydroxylysine aldehyde-initiated, or hydroxyallysine cross-linking pathway [18].

A segment of the allysine cross-linking pathway is shown in Figure 1.3. The diagram of allysine cross-linking pathway, shows the enzyme lysyl oxidase, which catalyses the oxidation of lysine at the telopeptide region of procollagen to form the telopeptide allysine (its aldehyde form), where the telopeptide region of procollagen can be defined as the two opposite extreme termini of the triple helix that remain unwound [31]. This telopeptide allysine residue reacts with the hydroxylysine residue of a nearby helix forming hydroxylysino-norleucine, a divalent cross-link often described as ‘immature’ [18, 19]. When the telopeptide allysine reacts with the nearby telopeptide allysine, it forms a dimer, which is also a divalent cross-link. In cases where histidine participates in cross-link formation, mature cross-links are formed. These mature cross-links are multivalent in nature [18].

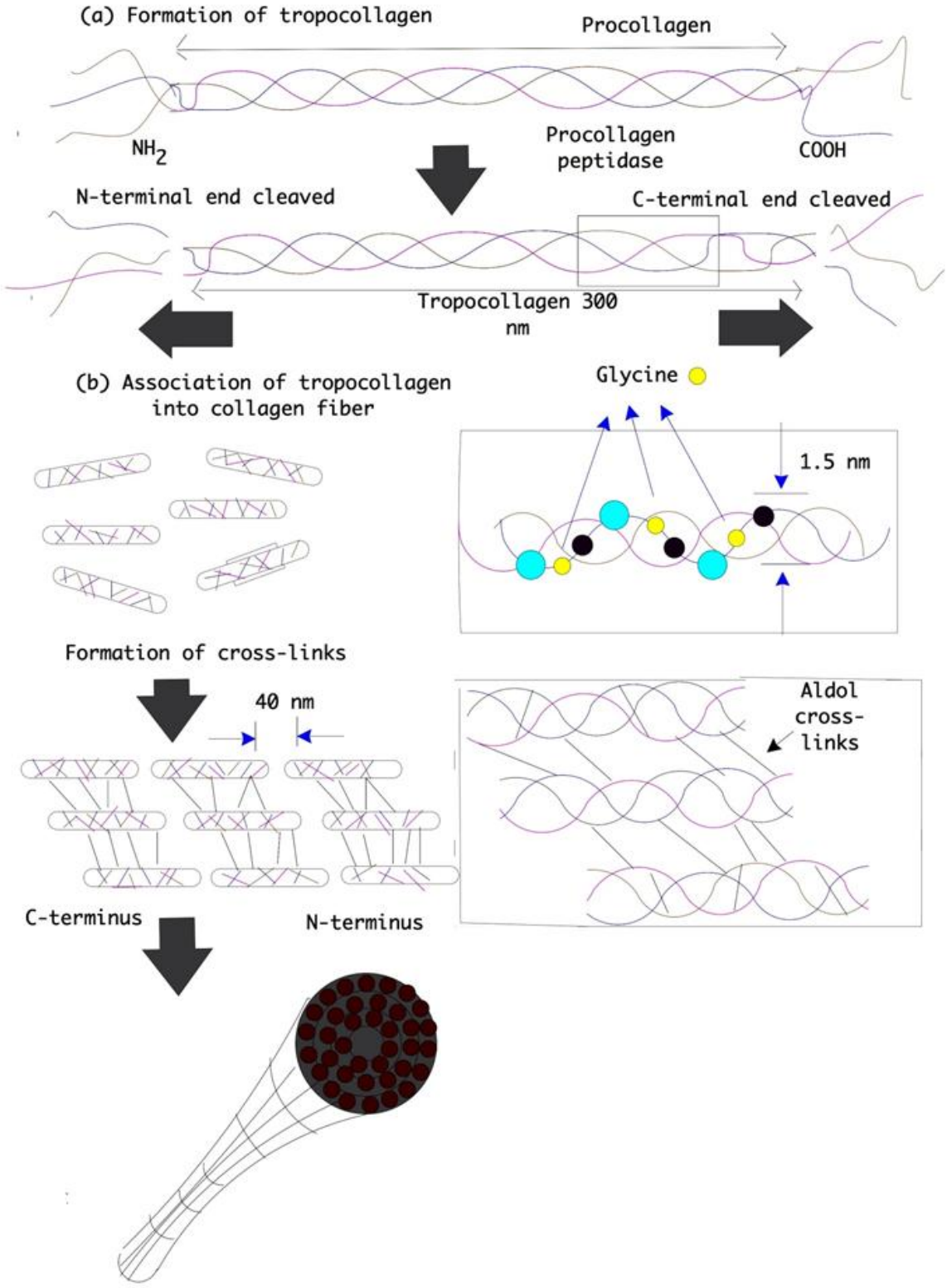


Figure 1.2 General process of collagen formation. Figure adapted from [39, 40] and drawn in ConceptDrawPRO.

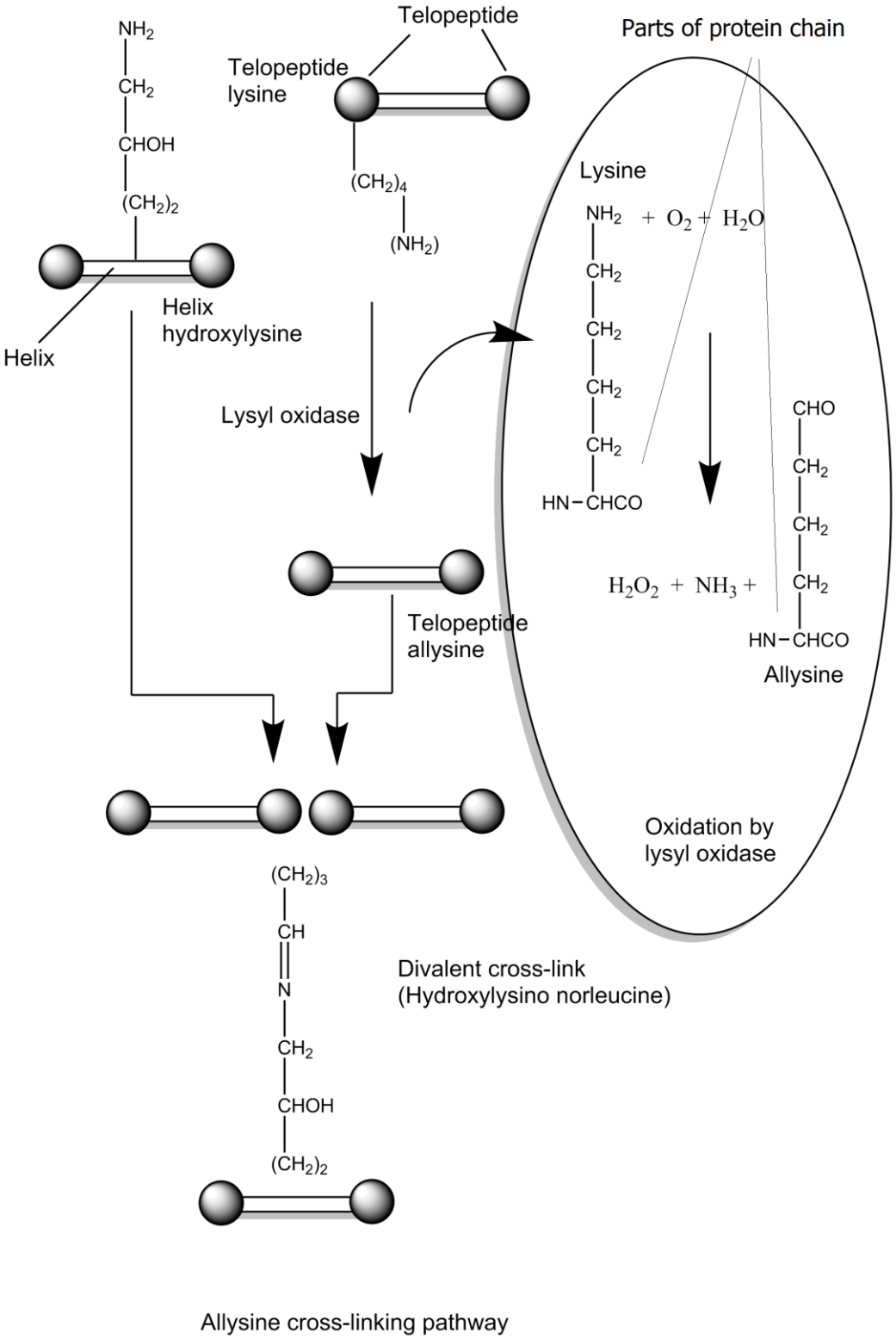


Figure 1.3 Allysine cross-linking pathway. Figure adapted from [41] and drawn in Chemdraw.

A segment of hydroxyallysine cross-linking pathway is shown in Figure 1.4

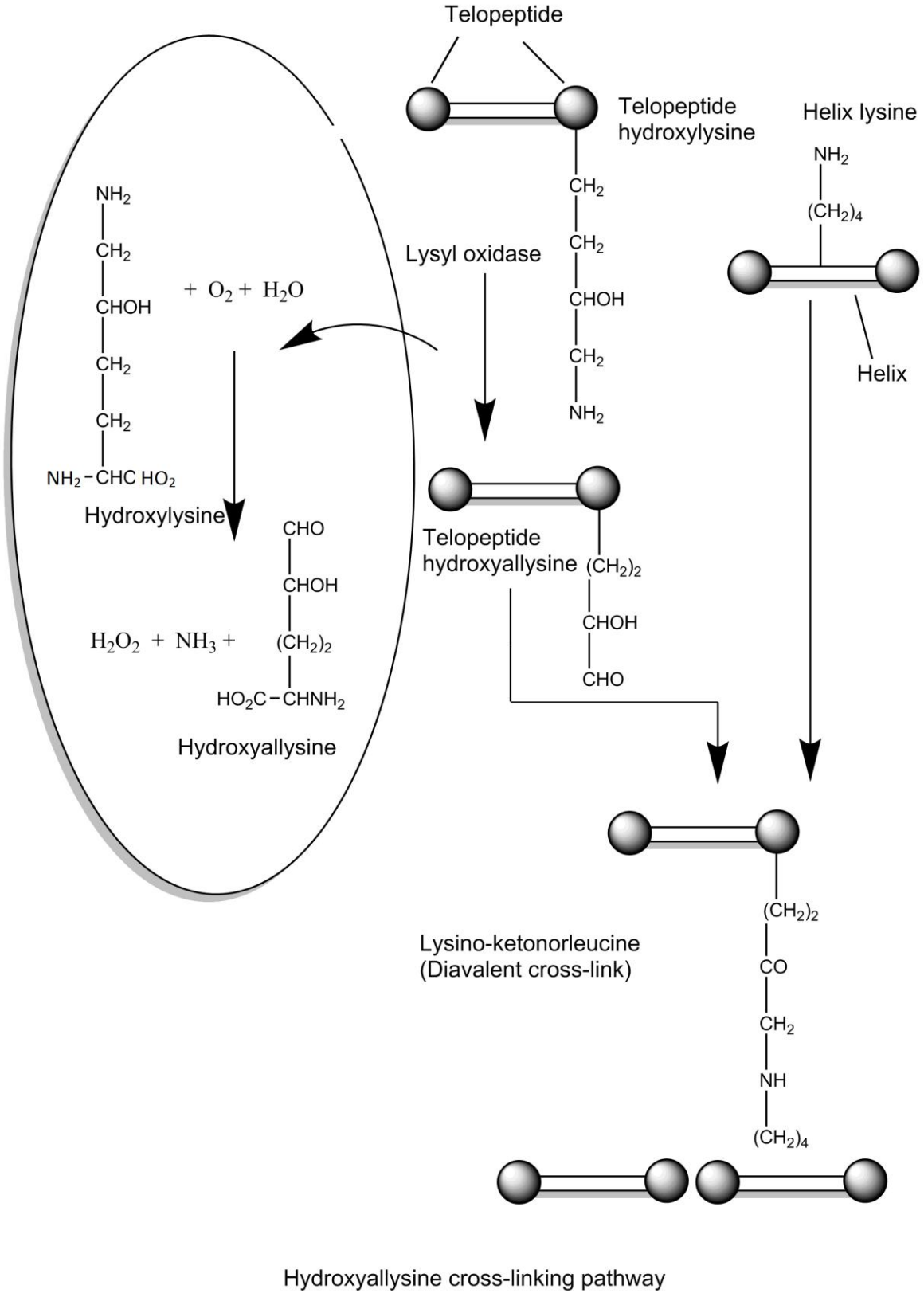


Figure 1.4 The hydroxyallysine cross-linking pathway. Figure adapted from [41] and drawn in Chemdraw.

Hydroxylysine in the telopeptide portion of procollagen is oxidised by lysyl oxidase to form the telopeptide allysine (aldehyde form) in Figure 1.4, the hydroxyallysine cross-linking pathway. This reacts with the lysine residue from the nearby helix to form lysino ketonorleucine, the divalent cross-link [18]. Effective addition of another hydroxyallysine forms a trivalent or 'mature' cross-link [18, 21].

These cross-links are shown in Figure 1.3 and Figure 1.4; they are affected by growth rate of the animals, nutrition and genetics [2]. In younger animals, the collagen cross links are reducible, whereas in older animals they are comparatively mature, thermally stable and less soluble [2, 25]. Reducible cross-links are the non mature divalent cross-links [42].

1.3.3 Mimicking the structure of collagen

The small systems (collagen mimetic peptides) are well suited for high resolution structural techniques like X-ray crystallography and Nuclear Magnetic Resonance (NMR) and have resulted in the first high resolution collagen triple-helical structure [32].

Scientists have chosen to study these 'collagen model peptides (CMP)' because of the capacity to mimic 'native collagen's' ability of forming linear assemblies. This ability has been a milestone for designing 2D or 3D fiber networks containing chemically, biologically or optically active elements [43]. Nanostructures formed by the self-assembly of CMPs and peptide amphiphiles have provided important applications in bio-engineering including scaffolds for bone regeneration, differentiation of stem cells or tissue scaffolding and diverse applications in bio-nanotechnology and regenerative medicine [44].

CMPs can self assemble into stable homotrimeric and heterotrimeric triple helices [32] (see Figure 1.5). The self-assembling properties of the peptide sequences can be assessed by two methods: (1) solution studies [45, 46] and (2) crystal structure studies [47].

Solution studies include turbidity analysis [48], Dynamic Light Scattering (DLS), Differential Scanning Calorimetry (DSC), Analytical Ultra Centrifugation (AUC), Circular Dichroism (CD) [48], Transmission Electron Microscopy (TEM), [48] Scanning Electron Microscopy (SEM) and Atomic Force Microscopy (AFM) [48]. Crystal studies involve analysis by X-ray fiber diffraction and X-ray crystallography methods [47].

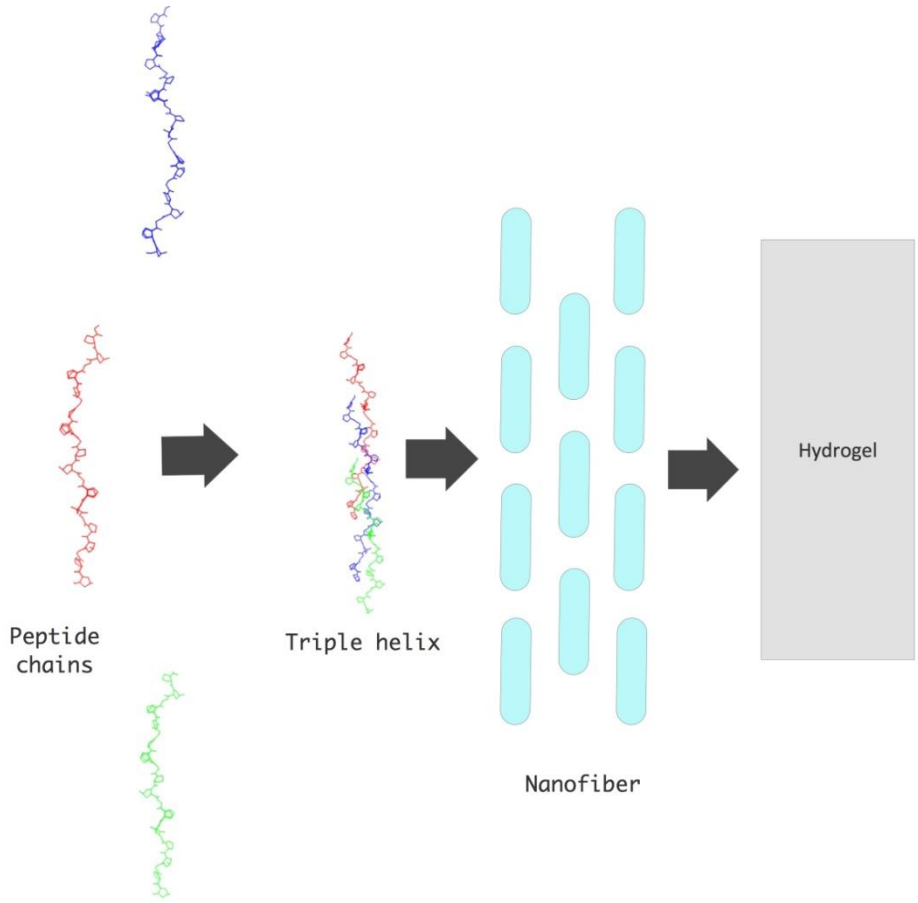


Figure 1.5 Formation of higher order structures, nanofiber, then a hydrogel. Figure adapted from [48]. Figure drawn in ConceptDraw PRO.

1.3.4 Homotrimers

1.3.5 Structure of triple-helix forming homotrimers

Simple peptide motifs can be engineered using the structure of natural collagen to form triple-helix forming sequences, which assembles into homotrimers (see Table 1.1). In solution studies, the peptides tend to precipitate out in the form of higher order structures [48].

Table 1.1 Triple-helix forming homotrimers

No.	Peptide sequence	Number of residues	Higher order assembly	Conditions	References
Design 1 - Simple trimeric motifs					
-1	$(PPG)_{10}$	30	no	Crystal	Okuyama et al. 1981 (Okuyama,
			no	Crystal	(Kramer, Vitagliano et al. 1998)
-2	$(POG)_{10}$	30		Crystal	Long et al. 1993 (Long, Braswell et al. 1993)
			no	Solution	Nagarajan et al. 1999 (Nagarajan, Kamitori et al. 1999)
			branched fibrils		
				Crystal	Li et al. 1993
			Branched fibrils	Solution	(Kar, Amin et al. 2006)

-3	$(OPG)_{10}$	30	No	Solution	(Inouye, Kobayashi et al. 1982)
-4	$(P\text{-fluoroP-G})_{10}$	30	no	Solution	(Holmgren, Bretscher et al. 1999)
Design 2- N- and C-terminal appended trimeric motifs					
-5	F-$(GPO)_{10}$-F	32	microfibrils	Solution	Cejas et al. 2008 (Cejas, Kinney et al. 2008)
-6	pentafluoroF-$(GPO)_{10}$-F	32	microfibrils	Solution	Cejas et al. 2008
-7	L-$(GPO)_{10}$-F	32	no	Solution	Cejas et al. 2008
-8	H-$(GPO)_{10}$-H	32	no	Solution	Cejas et al. 2008
-9	pentafluoroF-$(GPO)_5$-F	17	no	Solution	Cejas et al. 2008
-10	$(POG)_9$-POA	30		Crystal	Bella et al. 1994 (Bella, Eaton et al. 1994)
Design 3- N- and C-terminal simple trimeric flanking motifs					

					Bella, et al. 1995, 1996
-11	$(POG)_4$-POA-$(POG)_5$	30		Crystal	(Bella, Brodsky et al. 1995, Bella and Berman 1996)
-12	$(POG)_4$-QKG-$(POG)_5$	30		Crystal	(Kramer, Venugopal et al. 2000)
-13	Ac-$(GPO)_3$-GPR-$(GPO)_4$-GG-NH_2	26	no	Solution	(Yang, Chan et al. 1997)
-14	Ac-$(GPO)_2$-GFOGER-$(GPO)_3$-NH_2	21		Crystal	(Emsley, Knight et al. 2000)
-15	$(POG)_3$ITGARGLAG$(POG)_4$ (T3-785)	30		Crystal	(Kramer, Bella et al. 1999)
					(Li, Fan et al. 1993)
				Solution	(Liu, Siegel et al. 1996)

	$(GPA)_2GPVGPAGARGPA(GPO)_4GV$				Buevich et al. 2000
-16	(T1-892)	30	no	Solution	(Buevich, Dai et al. 2000)
			no	Solution	(Xu, Bhate et al. 2002)
Design 4- N- and C-terminal "sticky ends" trimeric flanking motifs					
-17	$(PKG)_4$-$(POG)_4$-$(DOG)_4$	36	fibers, gel	Solution	(O'Leary, Fallas et al. 2011)
-18	$(PRG)_4$-$(POG)_4$-$(EOG)_4$	36	microfibers	Solution	(Rele, Song et al. 2007)
Others					
-19	TREN-[suc-(Gly-Nleu-Pro)$_n$ -NH_2]$_3$ (n = 5 and 6)	45-54	no	Solution	(Kwak, De Capua et al. 2002)
-20	Boc-Ala-TRIS-[(G-Nleu-P)$_n$-OMe] triple helix in H_2O for n = 5, 6, 8 not for n = 1, 4	16-25	no	Solution	(Cai, Wong et al. 2007)

Of the 20 motifs reported to date, 17 consist of at least 30 residues and three consist of a minimum of 20 residues [32], see Table 1.1.

Higher order assemblies have only been observed in original POG triad rather than PPG, suggesting a specific role of proline oxidation in the superstructure of collagen. $(POG)_{10}$ has demonstrated self-association into nanoscale fibrils, while $(PPG)_{10}$ was reported to undergo amorphous aggregation (precipitation) under similar conditions [49]. This result supports a specific role of proline oxidation in the association of collagen triple-helices into higher order fibers.

Oxidative modification is quite common in meat and meat-derived products, due to the processing, environmental or consumer-induced oxidative insult [15], such as cooking. Oxidative modification may have direct relevance in understanding the impact of processing, handling and food preparation on meat quality. This hypothesis has been tested, in the last chapter of this thesis work with collagen model peptides.

In 2006, Brodsky et al. reported a high order structure for $(POG)_{10}$ peptides, in which multiple triple helices aggregated together to form mesh-like assemblies but did not form an organized structure like fibers or hydrogels, Figure 1.6 [49]. It is noted that the inability of these CMPs to form fibrils and fibers could be because of the absence of "sticky ends". The sticky end peptide sequences have oppositely charged hydrophilic residues at the ends (i.e. arginine and glutamic acid residues) of the peptide sequence, which in turn help in binding the adjacent polypeptide chains and thus form fibrils. The sticky end CMPs have also been used in α-helical coiled coil design to drive the fiber formation [50]. Progressive higher order assemblies have been reported from sticky end motifs in literature.

There have been various contradictory studies regarding the impact of varying GPO chain lengths on the stability of triple helices. Suto et al, confirmed that the melting points (MP) varied with varying chain lengths of GPO peptides [51]. Han et al mentioned that the MP of helices remained similar with the varying chain lengths [1]. In this thesis, the effect of varying (GPO) peptide chain lengths on the properties of self assembly are studied.

The above discussed peptides are based on bottom-up design. A bottom up design refers to bringing together the smaller sub systems to form an organized working system [52].

1.3.6 Higher order assembly of homotrimers

From table 1.1 it is seen that although 20 different homotrimers form triple-helices, only five of these underwent further association into higher-order fibrillar assemblies in solution (Table 1.1) [32, 53]. Details below-

1.3.6.1 Branched fibrils of $(POG)_{10}$

The hydrated triple helical form of $(POG)_{10}$ has been reported to form a loose intermediate that either unfolds into denatured monomers or self-associates into reversible branched nanometer-scale fibrils, see Figure 1.6 [49]. It seems that fibril formation was supported by neutral pH and promoted by temperature [48]. The presence of sugar was shown to disfavor fibril formation, likely due to the interruption of hydrogen-bonds between the triple-helices [54].

1.3.6.2 Formation of micofibrils by C-terminal appended trimeric peptide motifs

The incorporation of two aromatic residues (FF or pentafluoroF-F) at the C-terminal end of $(GPO)_{10}$ results in the formation of micrometer-scale fibrils based on triple-helices. Incorporating hydrogen or phenylalanine at both ends of the peptide, does not induce fibril formation [55].

Based on these reports, it is understood that assembly properties are influenced by incorporating changes at the terminals. Hence in this research work, the NH_2-$(GPO)_7$-NH_2 peptide have been investigated by replacing the usual C terminal of NH_2-$(GPO)_7$-COOH collagen model peptide, with an N terminal and comparing both of their assembly properties. NH_2-$(GPO)_7$-NH_2 was initially ordered due to the ease of its synthesis.

From previous studies of Cejas et al., $(GPO)_{10}$ peptides, with pentafluro-phenylalanine at the N-terminal and phenylalanine at the C-terminal, fibrils were formed [56]. DLS could not differentiate between the peptides. AFM has proved useful for imaging $(GPO)_{10}$ with phenylalanine at both ends. These readily assembled into micrometer scale fibrils, having stable helices. For $(GPO)_{10}$ with leucine and phenylalanine at its ends for hydrophobicity, no assembly was seen. $(GPO)_{10}$ with hydrogen at its end did not self-associate as readily [56].

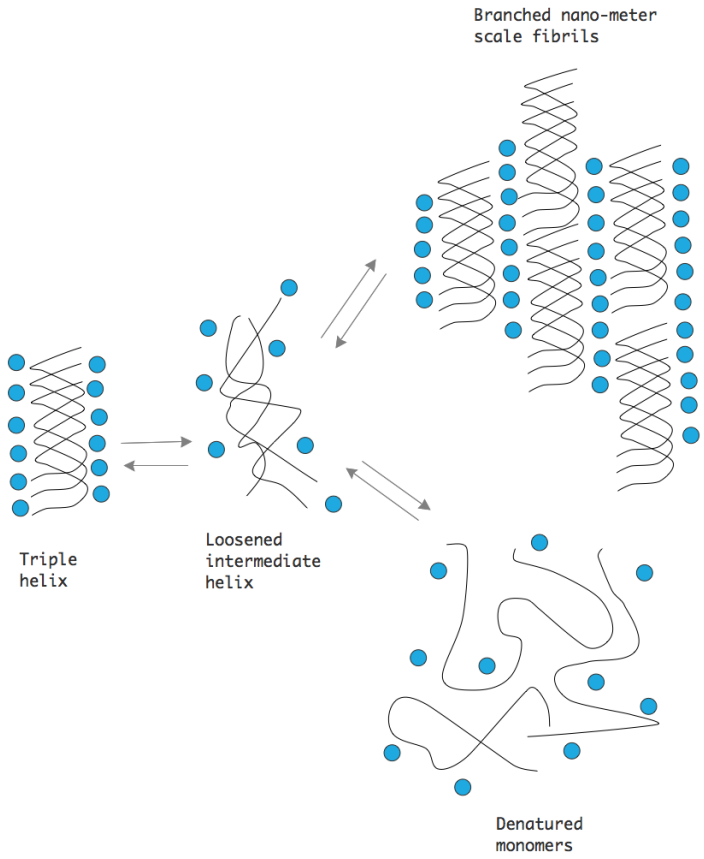


Figure 1.6 Self association process of (Pro-Hyp-Gly)$_{10}$. Figure adapted from [49] and drawn in ConceptDraw PRO

1.3.6.3 Nanofibers by motifs with N- and C- terminal "sticky ends" flanking motifs

From previous literature, Song et al have studied the fibrillar higher ordered structure formations by collagen mimetic peptides where they found electrostatic interactions as the "sticky ends" flanking motifs within the homotrimeric designs [57].

1.4 Residue modifications research in this thesis

Oxidative modifications can be introduced into muscle proteins at any stage post-slaughter – during meat formation, processing, storage and transport, retail, and preparation for consumption. There is limited knowledge about the effect of the progressive modification on the physico-mechanical properties of meat and meat-derived products.

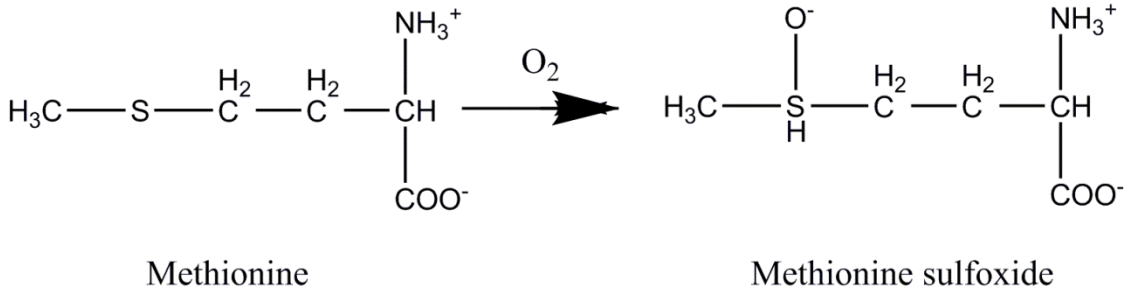
Targeted modification of CMPs will facilitate investigation of how specific oxidative modifications affect subsequent supramolecular assembly of collagen peptides. Since cooking processes are fundamental components in the processing and consumption of meat products, studying modifications related to hydrothermal insult is the focus here. On heat treatment, the amino acid residues in proteins are modified, with these changes at the primary structural

level progressively affecting higher orders of structure. Residue-specific side-chain modification, backbone cleavage, crosslinking, carbonylation and peptide terminal modifications all influence the secondary, tertiary and quaternary structure of the protein (in this case, collagen) [14].

Our hypothesis is - the oxidation events within the protein primary structure will result in changes to the self-association properties of the protein and its derivative peptides, which in turn may prove to be of particular importance in the manufacture of reconstituted foods from meat proteins.

During meat processing, e.g. cooking, there is production of reactive oxygen species (ROS), and these are highly likely to play a significant role in modification of meat properties associated with cooking. Oxidative modification is induced through initial generation and attack of ROS, or by reaction with secondary by-products of oxidative stress, such as hydroperoxides. The amino acids most susceptible to oxidative attack are cysteine and methionine, e.g. in case of cysteine, oxidation leads to the formation of disulfide bonds, but also to irreversible damage such as the formation of cysteic acid [15, 58]. Other products of oxidative modifications (see Figure 1.7), include: glutathiolation; methionine sulfoxide; carbonyl formation in amino acids such as lysine, arginine, proline and threonine; oxo-histidine; dityrosine, chlorotyrosine, nitrotyrosine; kynurenine derivatives of tryptophan; chloramines, deamination, and deamidation. Other modifications that can occur when metals catalyse the oxidation of amino acid residues include: formation of aspartate, asparagine, oxo-histidine from histidine; formation of hydroxyproline, glutamate, glutamylsemialdehyde from proline; glutamylsemialdehyde from arginine; amino-adipicsemialdehyde from lysine; amino-adipicsemialdehyde from threonine; dityrosine from tyrosine and di-sulfide cross-links from cysteine [59]. One notable form of oxidative modification in relation to collagen is hydroxylation, which can affect a wide range of amino acid residues, but particularly the aromatic and cyclic residues, including proline.

It has been seen that the presence of hydroxyproline enhances the self assembly of peptide chains [60]. When $(PPG)_{10}$ and $(POG)_{10}$ were introduced to self association conditions, the CMP without hydroxyproline did not undergo fibril formation whereas CMP with hydroxyproline supports self association, forming a hydration network (see Table 1.1).



Methionine Methionine sulfoxide

Figure 1.7 Example of oxidative modifications: Methionine oxidation. Figure made in Chemdraw.

This network involves intermolecular hydrogen bonding interactions between triple helices [48, 61, 62]. It has been also been noted that the specific cooking parameters utilized will likely have a significant effect on amino acid residue modification, for instance hydrothermal versus flame cooking [8, 14-16]. This research focuses on modifications likely to occur in proteins during cooking and their impact on collagen self assembly.

1.5 Heterotrimers

Most published work have used homotrimeric collagen mimics (AAA) involving collagen type II and type III found in cartilage and skin. Self assembly of the heterotrimers, AAB/ABC still remains a scarcely studied subject [32].

1.5.1 Assembly of heterotrimers

1.5.1.1 AAB motifs

The first example of a self assembled heterotrimer was described by Brodsky et al. in 2009, who synthesized a peptide system with POG flanking regions and sequences from α1 and α2 chains of type IV collagen that formed an AAB type collagen triple helix, which was analyzed by CD, DSC and NMR [48], (see Table 1.2).

Raines and co-workers formed a self-assembled heterotrimer from $(PPG)_7$ and a peptide containing fluoroproline derivatives at the X and Y positions, but it failed to fold into a stable triple helix due to unfavorable steric interactions [63].

Another approach towards self-assembly was that, oppositely charged amino acids were used to drive the self-assembly of stable heterotrimeric collagen mimetic peptide. Mixing three peptides, one with charge +n, one with -n and one neutral, a zwitter ionic triple

helix was formed. Using different values of n, CMPs with ten triplets were investigated where arginine and glutamic acid residues were used as charged moieties and the POG sequence as the neutral species [64]. Self-assembly was also achieved by utilizing favorable electrostatic interactions, which successfully produced an AAB heterotrimer with control over register and composition. When, by mixing two peptides with a charge -5 $(EOGPOG)_5$ and one with +10 charge $(PRG)_{10}$ a zwitterionic AAB triple helix was formed, it was found to be more stable than the ABC heterotrimer which has been previously studied [65].

Table 1.2 Triple-helix formation in the solution as heterotrimers; Higher order assemblies were not visualized

	Peptide sequence	Number of residues	References
Design 1 - AAB motifs			
-21	A=$(PPG)_{10}$ B=$(POG)_{10}$ and B=$(PPG)_{10}$ A=$(POG)_{10}$	30	(Slatter, Miles et al. 2003)
-22	**A**= $(EOGPOG)_5$ **B**= $(PRG)_{10}$	30	(Russell, Fallas et al. 2010)
-23	**A**= $(GPO)_3$GMOGVGEKGEOGKO$(GPO)_2$GY **B**=POGDO$(GPO)_2$G**ISL**KGEEGPOGP**A**GPOGYOG	31 32	(Madhan, Xiao et al. 2008)
Design 2 - AAB motif with cysteines			
-24	**A**=$(GPP)_5$GARGER**G**FP**G**ERGVQGPO$(GPP)_5$GAA **C** **B**=$(GPP)_5$GPR**G**ERGPOGESGAAGPT$(GPP)_5$G**CC**Y	52 52	(Slatter, Foley et al. 2006)
Design 3 - ABC motifs with cysteines			
-25	**A**= H_2N-$(GPO)_5$GPQGIAGQRGVVG**C**GG **B**= H_2N-$(GPO)_5$GPQGIAGQRGVVGL**C**GG	31 32	(Ottl,

			Battistuta et al. 1996)
	C= H_2N-$(GPO)_5$GPQGLLGAOGILG**CC**GG	32	
- 26	**A**= Ac-$(GPO)_3$GPOGDQGPOGIO$(GPO)_2$G**C**GG	31	Saccà et al. 2002 (Saccà, Renner et al. 2002)
	B= $(GPO)_3$GPOGDQGPOGIO$(GPO)_2$GP**C**GG	32	
	C= Ac-$(GPO)_3$GAKGRAGFOGLO$(GPO)_2$G**CC**GG	32	
- 27	$(POG)_{10}$ $(POG)_2$(**E**OGPOG)$_3$$(POG)_2$ (**E**OGPOG)$_5$ (P**R**GPOG)$_5$ $(POG)_2$(PRGPOG)$_3$$(POG)_2$ (Solution of the above 6 sequences formed random helices)		(Gauba and Hartgerink 2007)

1.5.1.2 ABC motifs with cysteine residue:

From previous studies, Fragments 1 ($(POG)_5$-$(POG)_3$-Cys-Gly) and 2 ($(POG)_3$-GlyCysCysGly-$(POG)_5$), were prepared in which the three helicogenic strands linked by a pair of disulfide bonds offset the strands and sets the register. $(ProYaaGly)_3$ segments in the two identical strands α1 and α1', with a third strand α2, formed an intramolecular triple helix. Intermolecular assembly of overhanging $(ProYaaGly)_5$ segments yield collagen triple helices [45]. Self assembly has also been performed on the covalently linked heterotrimers and self assembled heterotrimers, which lead to the formation of higher level self- assembled structures [48].

1.6 Comparative study of crystal structures of collagen peptides $(GPO)_{10}$ and $(GOO)_9$

It has been reported that $(GOO)_{10}$ forms stable triple helices [66] [67]. The literature reports that one of the factors, which supports the triple helical stability, was the proline puckering configuration of the proline residues. One of the factors that support the triple helical stability

was puckering configuration of amino acids. Usually 4(R)-hydroxyproline (HypR) relaxed in an up-puckered configuration, while proline had a property of taking either of the up or down puckered configuration [66], see Figure 1.8.

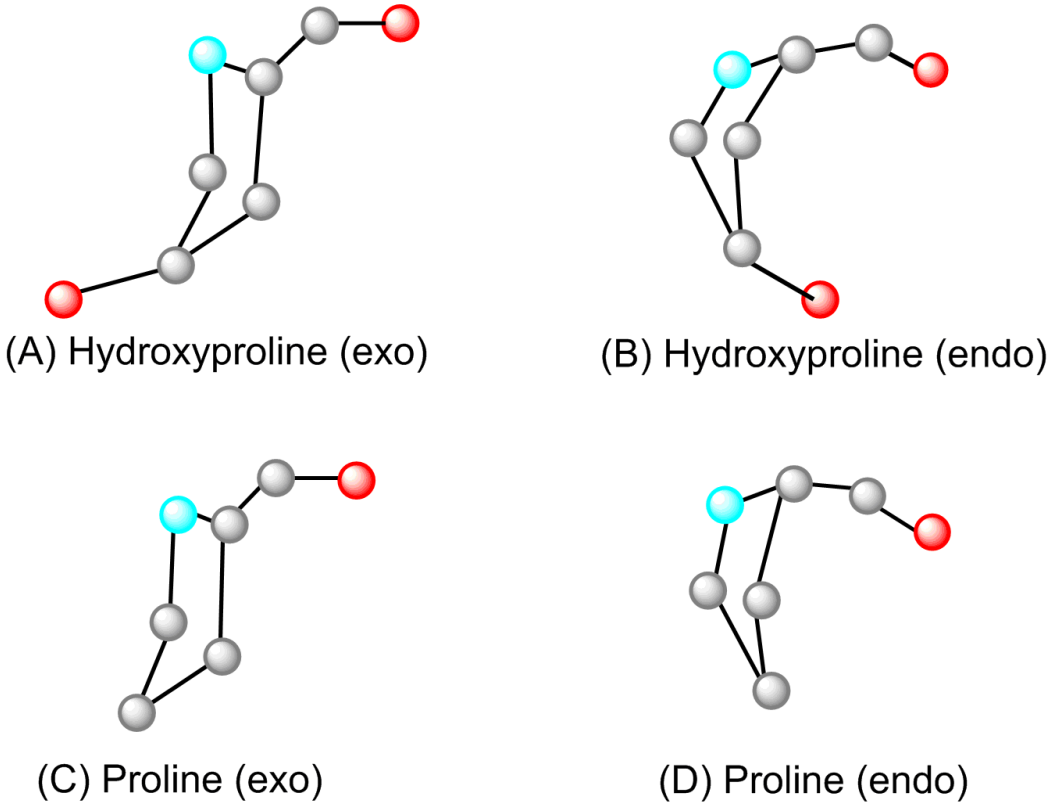


Figure 1.8 Proline puckering (A) Hydroxyproline (exo) up-puckered conformation (B) Hydroxyproline (endo) down puckered conformation (C) Proline (exo) up-puckered conformation (D) Proline (endo) down puckered conformation. Figure drawn in ConceptDraw PRO.

According to a propensity-based rule, it was required that the pyrrolidine ring pucker should be down at the X position and up at the Y position. But when $(OOG)_9$ was selected (where O is 4R hydroxyproline), it still formed a thermally stable triple helix, despite its contradiction to the propensity rule [66]. In $(OOG)_9$, (4R) Hyp residues in the Yaa positions took the Cγ exo position, the up-puckered conformation and the residues in the Xaa position took the Cγ up-puckered conformation too, unlike the other collagen triple helices.

It was observed that variations in triple helix pitch (the pitch of a helix is the width of one helix turn measured parallel to the axis of the helix) was proposed to play in important role in functional binding of collagen to other biomolecules, such as integrin or heparin [68].

While the helical pitch of $(GOO)_{10}$ was found to be 7/2 (3.5 residues per turn), consistent with the prediction for $(GPO)_{10}$, the natural collagen had a helical pitch of 10/3 [69]. Table 1.3 shows the difference in dihedral angles for helices of pitch 7/2 and pitch 10/3. The differences in torsion angles of $(GPO)_{10}$ and $(GOO)_9$ are shown as in Table 1.4.

Table 1.3 Idealized/ dihedral angles. 7/2 helical pitch stands for $(GPO)_{10}$ and $(GOO)_9$ while 10/3 is for collagen [70],[71]

	Pitch- 7/2	**Pitch- 10/3**
Gly Φ (^{o})	70.2	67.6
Gly Ψ (^{o})	175.4	151.4
Xaa Φ (^{o})	75.5	72.1
Xaa Ψ (^{o})	152.0	64.3
Yaa Φ (^{o})	62.6	75.0
Yaa Ψ (^{o})	147.2	155.8

The interchain H-bonding of $(GOO)_{10}$, $(GPO)_{10}$ and $(PPG)_{10}$ are also compared. It is observed that $(GPO)_{10}$ has shortest inter chain H-bonding, followed by $(GOO)_{10}$ followed by $(GPP)_{10}$ (see Table 1.5).

Table 1.4 Torsion angles for $(GPO)_{10}$ and $(GOO)_9$ [69]

	$(GOO)_9$ PDB- 1YM8	**$(GPO)_{10}$, PDB-1V7H**
	Resolution- 1.55 Å	Resolution – 1.24 Å
Gly Φ (^{o})	-73.1	-73.1
Gly Ψ (^{o})	175.6	174.2
Xaa Φ (^{o})	-70.5	-71.3
Xaa Ψ (^{o})	161.3	161.5
Yaa Φ (^{o})	-66.2	-56.9
Yaa Ψ (^{o})	156.4	150.0

Table 1.5 Interchain hydrogen bond parameters, Gly NH..OCXaa. Standard deviations are given in parentheses. Protein Data Bank (pdb ref. numbers are 1YM8, IK6F and 1V7H [66, 69]).

	$(GOO)_9$ 1YM8	**$(PPG)_{10}$ IK6F**	**$(POG)_{10}$ 1V7H**
Distance (Å)	2.92 (0.06)	2.97 (0.06)	2.90 (0.06)
Angle (^{o})	164 (3)	167 (2)	169 (4)

1.7 Thesis outline

In chapter 3 of this thesis, the effects of chain lengths of the (GPO) peptide sequences on the triple helical 1st order self-assembly formation and higher ordered assembly formation have studied. Biophysical structural characterizations of all the selected CMPs have been carried out. The differences in assembly properties of chain lengths of the shortest 9 residue peptide, $(GPO)_3$ and the longest 30 residue peptide, $(GPO)_{10}$ have been investigated. The medium chain length peptides, $(GPO)_5$, $(GPO)_7$, amidated $(GPO)_7$ and racemic $(GPO)_5$ sequences have been characterized. The effect of factors affecting triple helical self-assembly including temperature, pH, concentration of peptides, time and chain length have been studied. Triple helix formations and higher order structure formations were found to be directly proportional to chain length.

In chapter 4 of this thesis, an in-depth investigation of the impact of temperature cycle, on the assembly properties of (GPO) collagen peptides has been performed. Insights have been obtained using Infrared spectroscopy on crystals and heat-treated assemblies into the states of biomolecular interactions, especially disruption of H-bonds and polyproline II helical conformations at a primary level. Atomic force microscopy has been used to record the changes in assembly patterns, which occur from the temperature treatment follow-ups. Other analytical characterization techniques have been used to study protein chemistry, including CD, DSC, DLS, TEM, AFM, mass spectrometry (ESI and MALDI) and SAXS to support the findings in chapters 3 and 4.

In chapter 5, the atomic details of (GPO)n triple helical peptides by X-ray crystallography and X-ray fiber diffraction have been studied. The structure of a $(GPO)_{10}$ collagen triple helical peptide has been successfully determined. Extensive study has been

carried out on the intra and inter H-bonding network, the localization of residues within the hydrophobic core of helices, network of triple helical packing, proline puckering, helical pitch and symmetry, torsion angles and importance of water molecules surrounding the helix with the most up to date $(GPO)_{10}$ structure solved at of 0.89 Å resolution, in this chapter. Crystallizing of the racemic 16 residue peptide, $(GPO)_5$-Gly has also been successful. This has been the smallest possible collagen model peptide to form crystals, which diffracted at both 1 Å and 1.5 Å resolutions. Attempts with small molecular crystallography, protein crystallographic direct methods and molecular replacement methods have not provided a solution for the racemic peptide. Solution studies have been conducted on the L-$(GPO)_5$-Gly peptides.

In chapter 6, the modified peptide sequence $(GOO)_{10}$ has been introduced and characterized. $(GOO)_{10}$ is the oxidized version of $(GPO)_{10}$ and has the capability to give valuable insight into the differences occurring in assembly properties due to oxidation. In this peptide sequence, all the prolines have been hydroxylated. From previous literature and recent findings of $(GOO)_{10}$'s crystal structure, the higher stability of GOO was shown to be due to intra-strand water bridges, which adopt a stable monomer polyproline II conformation in the $(GOO)_{10}$ triple helix. However, this activity has not been proved enough by solution studies until today. With infrared spectroscopy on the GOO_{10} crystals, heat-treated GOO_{10} nano-assemblies, this hypothesis has been re-established (along with CD and DSC work). In the next part of this chapter, effect of the food processing technique, cooking has been investigated. The effect of hydrothermal insult, targeted at the primary structural level of the peptide sequence has been found instrumental in impacting on the supramolecular assembly properties. Its effect and impact on the supramolecular assembly properties have been studied. From a redox proteomic profiling approach, a trend has been observed such that, a range of recognized modifications were observed at the middle portions of the peptide, while the ends of the peptide chain were not affected. Tandem mass spectrometry was used to analyze amino acid modifications (MALDI-MS/MS, ESI-MS/MS, and/or ESI-$MS^{n\cdot}$). This trend has been established while studying the most common meat processing technique, cooking which further impacts on the texture of meat derived food products.

2 Materials and Methods

2.1 Peptides

Two kinds of collagen have been used in this research work (1) Bornstein and Traub calf skin collagen Type 1, C9791 SIGMA from Sigma Aldrich, (Auckland, New Zealand) and (2) Type 1 collagen from bovine achilles tendon. BSA protein was purchased from Sigma Aldrich. GPO peptides $(GPO)_3$, $(GPO)_5$, $(GPO)_7$, $(GPO)_{10}$ and $(GOO)_{10}$ were bought from Mimotopes, (Melbourne, Australia). Amidated $(GPO)_7$ and racemic peptides $(GPO)_5$-Gly were purified and provided by the Maurice Wilkins Centre (University of Auckland, New Zealand) in collaboration with Prof Margaret Brimble and Dr. Paul Harris.

Table 2.1 Table of peptides from their suppliers. At the N-terminus, H denotes free amine, and at the C-terminus, -OH denotes free acid. The sequences have been expanded below.

Peptides	**Company**
H-$(GPO)_3$-OH $(GPO)_3$-GPOGPOGOPO	Mimotopes
H-$(GPO)_5$-OH $(GPO)_5$-GPOGPOGPOGPOGPO	Mimotopes
H-$(GPO)_7$-OH $(GPO)_7$-GPOGPOGPOGPOGPOGPOGPO	Mimotopes
H-$(GPO)_{10}$-OH $(GPO)_{10}$-GPOGPOGPOGPOGPOGPOGPOGPOGPOGPO	Mimotopes
H-$(GOO)_{10}$-OH $(GOO)_{10}$-GOOGOOGOOGOOGOOGOOGOOGOOGOOGOO	Mimotopes
NH_2-$(GPO)_7$-NH_2 Also called amidated $(GPO)_7$	Maurice Wilkins Centre
Racemic mixes of L-$(GPO)_5$-Gly and D-$(GPO)_5$-Gly	Maurice Wilkins Centre

Mimotopes had assessed the purity for these (GPO) peptides by electrospray ionisation mass spectrometry and all peptides [$(GPO)_3$, $(GPO)_5$, $(GPO)_7$, $(GPO)_{10}$ and $(GOO)_{10}$] are >95% pure (Table 2.1).

Taking into account the high repeatability of specific residues in the collagen sequence, amino acids have been chosen and collagen model peptides have been designed for further work. Innovagen's peptide property calculator PepCalc.com was used to calculate the hydrophobicity and neutral charge of the chosen peptides. Expasy, Prot Param were used for calculating the % of amino acids present in sequences [72].

2.2 Reagents

Acetic acid, PBS (phosphate buffered saline), sodium chloride, sodium dihydrogen phosphate and disodium hydrogen phosphate were bought from Sigma Aldrich (Auckland, New Zealand). Trypsin and pepsin were purchased from Promega (Madison, USA). HCCA (α-Cyano-4-hydroxy-cinnamic acid) and calibrants were purchased from BRUKER (Bremen, Germany). Trifluroacetic acid, urea, sodium deoxycholate, ammonium bicarbonate, iodoacetamide (IAM), dithiothreitol (DTT) and other reagents have been purchased from Sigma Aldrich (Auckland, New Zealand). Polyethylene glycols (PEG) of molecular weights 300, 400, 550, 600, 800 and 2000 gram/moles were purchased from Hampton Research (California, USA). Acetonitrile, ethanol, methanol, tris-2-carboxyethyl phosphine 2 propanol (TCEP), dihydroxy benzoic acid (DHBC) and uranyl acetate were bought from Sigma Aldrich, (Auckland, New Zealand).

2.3 Materials

V-4 grade mica for atomic force microscopy was purchased from SPI. Plastic free safe lock Eppendorf tubes were purchased from Eppendorf (Hamburg, Germany). Plastic free tubes were used to avoid plasticisers in mass spectrometric analysis. Ninety-six welled three and four well sitting drop crystallisation plates, 24-welled hanging drop plates and plastic cover slides were purchased from Hampton Research (California, USA) and TTP labtech (MA, USA). Crystallisation screens for peptides were purchased from Hampton Research. C18, which is carbon 18 ZipTip stage tips were bought from Millipore Corporation, (Bedford, Massachusetts) for desalting of samples before mass spectrometry.

2.4 Instruments

pH measurements were carried out on an Ultra Basic UB 10 pH meter (Denver Instrument Co., Colorado, USA) fitted with a high performance glass body pH/Tris electrode. For centrifugation of small volumes (<2 ml), an Eppendorf 5810R with a fixed F-45-30-11 rotor was used, whereas larger volumes were centrifuged using a Thermo Scientific Sorvall RC6 plus centrifuge. For circular dichroism (CD) spectroscopy, cuvettes were bought from

JASCO. The Nanodrop spectrophotometer was used to determine concentrations. This was a Thermo Scientific NanoDrop™ ND-1000 Spectrophotometer (Thermo Fischer Scientific)

2.5 Buffer exchange

2.5.1 Size exclusion chromatography

A Superdex peptide column 10/300GL has been used for gel filtration chromatography. The flow rate was maintained at 0.75 ml/min. The column volume was 24 ml, while the capacity of the sample volume loop was 500 µl. Thus the sample volume required for injection was 600 µl.

For preparation of buffers, the buffer was filtered with a 0.2 µm filter and sonicated for 50 minutes, with the lid loose.

For sample preparation, samples were spun at 13,000 rpm at 4 °C for 5 minutes. The readings in mAU (absorption units) were taken at 215 nm and 280 nm wavelengths, but as the selected collagen model peptides have no aromatic amino acids present in them, only readings at 215 nm were considered. Absorbance at 215 nm shows the presence of peptide bonds.

2.6 Collagen solubilization

2.6.1 Solubilization

The procedure was adapted from AgResearch's in-house protocols. Bornstein and Traub calfskin collagen Type 1 (C9791 SIGMA) was soluble at 1 mg/ml in 0.1 M acetic acid when stirred at room temperature for 1-3 hours until dissolved.

2.6.2 Digestion

The method of collagen digestion was adopted and modified from [73]. 10 ml of 5 mM dithiothreitol (DTT) in 25 mM ammonium bicarbonate at pH 7.4 was prepared. Type-1 collagen, 260 µg was weighed and added to 100 µl of 5 mM DTT, 25 mM ammonium bicarbonate solution at pH 7.4 in a 0.6 ml Eppendorf. It was sonicated in a water bath at a temperature of 4 °C and 1.2 µg of trypsin was added. After sonication for 1 minute, the reaction was stopped for 30 seconds and sonicated for another 1 minute. Iodoacetamide was added at a concentration of 15 mM. The mixture was then incubated in dark for 45 minutes. Following incubation, 1 % of surfactant sodium deoxycholate (0.002 g) was added to the

solution followed by 1.2-µg trypsin (12 µl of 0.1 µg/µl stock) and the mixture was sonicated. It was later placed in a 37 °C water bath and incubated overnight.

2.6.3 Matrix Assisted Laser Desorption/Ionization (MALDI)

2.6.3.1.1 Matrix for peptides

The protocol for matrix preparation was adopted from the MALDI-TOF MS instrument manual from Bruker Daltonik, (Germany). HCCA was added in 75 % acetonitrile followed by 0.1 % trifluroacetic acid (TFA) to form matrix. When peptides are used as the analytical sample, di-hydroxy benzoic acid (DHBC) is used as the matrix, but for protein samples such as collagen, sinnapinic acid is usually used as the matrix.

2.6.3.1.2 Preparation

Following sample preparation method in the instruction manual for Bruker, C18 stages were washed with acetonitrile. Once acetonitrile bonds to the stage tips, pre-conditioning or equilibration is carried out with around 10 µl of 0.5 % of formic acid. The proteins/peptides were loaded with a syringe that increases the air pressure on the stage tips and hence binding occurs at the filter. The syringe is washed again with 0.5% formic acid before loading another protein/peptide. This protein/peptide solution is then mixed with a matrix and the mixture is placed in the ground steel/anchor chip target plate.

2.7 Collagen solubilization (interface layer of proteins)

AgResearch's in-house protocol has been used for collagen solubilization. By adding 12.1 g of tris.HCl in 200 ml of water, tris stock was prepared. The pH was adjusted to 7.8 with 2 M and 8 M HCl. Water was added until the concentration became 0.4 M. The denatured solution has been prepared with 6 M urea and 100 mM Tris buffer, i.e. 0.4 g of urea was added in 250 µl of Tris stock and vortexed. The volume was made up to 2 ml with water. Reducing agent has been made by mixing 30 mg of DTT in 750 µl of water, to which 250 µl of Tris stock was added and the resulting solution was vortexed. The final concentration was 200 mM DTT and 100 mM Tris. This solution was kept in the shaker for an hour to check if the collagen had solubilized. If collagen still does not dissolve, the solution is left for overnight incubation. If collagen still does not dissolve, the mixture is spun at 14,000 rpm for 2 minutes and 100 µl of the supernatant is taken out. Urea is added to give a final concentration of 8 M. After 1 hour, the mixture was sonicated for 5 minutes and incubated for 1 hour at room temperature.

Protein precipitation was performed as follows: to 100 µl of collagen solution, 800 µl of methanol is added and the mixture is vortexed. To this solution, 200 µl of chloroform was added, vortexed and then 600 µl of water is added. It was then spun for 1 minute at 14,000 rpm. The aqueous supernatant layer is removed and 800 µl of methanol is added and spun at 14,000 for 2 minutes. Methanol is then removed without disturbing the pellet. Protein is taken out from the interface layer and the pellet is air-dried.

After protein precipitation, 100 µl of protein solution (containing interface layer proteins) is taken out. To this solution, 50 mM ammonium bicarbonate is added and vortexed. 100 µl of TCEP is added to the above solution. This solution is placed in a hot water bath at 56 °C for 45 minutes. After 45 minutes, 10 µl of 50 mM IAM prepared in 100 mM ammonium bicarbonate is added and the resulting mixture is placed in a shaker for the next 30 minutes. To this solution, 2 µl of trypsin (l µg/ml Promega) is added and the solution is placed in a water bath for the next 18 hours at 37 °C.

2.8 Protein and peptide characterization

2.8.1 SDS-PAGE

The protocols for running SDS-PAGE of collagen and collagen peptide samples has been taken and modified from [74, 75]. SDS-PAGE analysis has been carried out using a Novex Sharp pre-stained protein ladder of protein molecular mass markers on gels, which gave 12 bands marking masses from 3.5 to 260 kDa. Bio-Rad pre-stained protein ladder has also been also routinely used, which gave 12 bands marking masses from 2 to 250 kDa. For SDS PAGE, 5 µl Novex® Sharp protein standard (Invitrogen) has been used as a protein ladder.

Gels were obtained from Invitrogen. They were composed of a resolving gel to separate proteins by molecular masses. NuPAGE Novex® 4 – 12 % Bis-Tris gels or 4 - 20 % Tris-glycine gels have been run according to the manufacturer's recommendations in a NuPAGE® gel electrophoresis box at room temperature. Novex gels of 4 to 20 % tricine gel have been used. A ten or fifteen-toothed comb was placed into the top of the gel to create wells. The gels were stained with Coomassie blue staining solution. Gels were de-stained for at least one hour. To facilitate staining and de-staining, the gels were briefly heated in a

microwave oven for 30 seconds to approximately 50°C. Gel photographs were taken using a Genius2 Bio Imaging System (Syngene).

2.8.2 Mass Spectrometry

Both MALDI and electrospray techniques have been used to check for oxidations on samples of collagen and collagen mimetic peptides.

2.8.2.1 Thin layer affinity HCCA Anchor Chip preparation.

For MALDI, the thin layer affinity HCCA anchor chip preparation procedure has been adopted from protocols of Goborn et al. [76]. An aqueous droplet containing the analyte was used and contaminants (e.g. salts) were deposited on the anchor covered with HCCA thin layer. The analyte molecules attach to the HCCA thin layer whereas water-soluble contaminants stay in a liquid phase and get removed together with the residual droplet after a few minutes.

2.8.2.2 Preparation protocol for thin layer affinity HCCA Anchor Chip

The protocol for HCCA anchor chip preparation has been adopted from Bruker's manual. Matrix solution is prepared by mixing acetone and 0.1 % TFA in the ratio 97:3. This solution is later saturated with HCCA crystals. It was then sonicated, centrifuged and supernatant was taken out.

The re-crystallization mix has been prepared as follows: HCCA is weighed, 0.2- 03 mg (a few HCCA crystals held in pipette tip) is added to a solution composed of ethanol: acetone: 0.1 % TFA in the ratio of 6:3:1. The sample analyte solution is then prepared that comprises of proteins and peptides dissolved in their respective buffers. The sample does not contain any organic solvents (which would dissolve the HCCA layer). The wash buffer was either 0.1 % TFA or 10 mM diammonium citrate (in 0.1 % TFA).

2.8.2.3 Procedure of thin layer affinity HCCA Anchor Chip

Two microliters of matrix solution was applied on a 600 µm anchor and removed with a pipette immediately. Within a second, the HCCA thin layer is formed. 1-4 µl of aqueous analyte solution is deposited on this thin layer and incubated for around 3 minutes; care is taken that the sample does not dry out. 2-4 µl of the washing buffer is added to the residual liquid and the whole droplet is removed completely as possible. It is observed that, usually the anchor retains 10-100 nl. Addition of 0.1% TFA before removing the analyte droplet dilutes the contaminants. 1 µl of re-crystallization mix is later added which dissolves the thin

layer with attached analyte molecules in it and adds some more HCCA to form larger crystals which would allow more laser shots per shot in MS [77].

2.8.2.4 Data analysis for mass spectrometry

Spectra have been evaluated using the mass spectrometric software Protein scape from Bruker. Initial peptide modification interpretation and searches for oxidations have been conducted [78]. Area under the oxidised peak has been accounted for. By dividing the area of base peak from the oxidised peak area, respective oxidation percentages have been calculated.

2.8.3 Biophysical characterisation techniques

2.8.3.1 Circular Dichroism (CD)

CD spectroscopy on collagen model peptides has been performed using protocols from Engel et al. and Bachinger et al. [48, 79]. CD data sets have been collected on Jasco J-815 circular dichroism spectrophotometer. Quartz cuvettes of path length 1 mm have been used. Initially protein and peptide solutions of concentrations 0.2, 0.5 and 1 mg/ml were prepared in various buffers and in water for analysis. The concentration screening was performed to check for the minimum concentration, which would form secondary structures to be examined by CD. The CD spectra analysis manager software has been used to set the parameters and take readings. Automatic scanning and manual scanning for temperature studies have been carried out at a scanning rate of 1°C/minute. The parameters used for (I) spectrum measurement (II) temperature measurement are as follows:

2.8.3.1.1 Spectrum measurement for CD

Spectrum measurement has been performed at a standard of 100 mdeg, while the wavelength scan was set up at a default scan of 260 nm to 170 nm. A data pitch was set up at 2 nm. Scanning mode was set up at a continuous speed of 200 nm/sec. The response was set up at 1-second intervals, bandwidth at 1 nm and accumulations were set up at 3 for each measurement.

2.8.3.1.2 Temperature measurement for CD

Manual wavelength has been set up at 225 nm. A positive maximum is the characteristic of triple helix formation at 225 nm. The start temperature was fixed at 20 °C and the end temperature was set up at 90 °C. The data pitch was 1 °C and the temperature slope was

maintained at 1 °C/min. The sensitivity has been 100 mdeg; the response was set at 1 sec and the bandwidth was set at 1 nm.

A wavelength scan has been carried out at 20 °C on all GPO peptides. A maximum around 222 nm – 225 nm indicates the presence of triple helix [48] . A minimum at 200 nm shows the presence of random coils [80].

2.8.3.1.3 Data analysis for CD

The graphs were plotted such that the ellipticity generated from the Jasco CD spectra instrument's software was then converted to mean residual ellipticity (MRE). MRE was plotted in the Y-axis and wavelength in the X-axis. The formula for ellipticity conversion is described elsewhere.

2.8.3.2 Dynamic light scattering (DLS)

The procedure for DLS analysis of collagen model peptides has been taken and modified from Kar et al. and Brodsky et al. [48, 81]. The equipment was operated according to the Malvern instrument supplement manual. This has been used to analyse the size distribution profile and hydrodynamic radius of particles in solution. Square polystyrene cuvettes (DTS0012) of 12 mm, from Malvern instruments, (Malvern) UK were used. DLS relies on the measurement of the light scattering interference caused by particles, which move due to Brownian motion. The interference therefore depends on the size of the particles since larger particles move slowly whereas small particles move quickly through the solution. To avoid dust and aggregation, samples were prepared carefully. Dynamic light scattering has been performed on a Malvern Zetasizer Nano (Malvern, United Kingdom). For analysis, the Malvern Zetasizer software has been used. Peptides were chosen as the scattering material and water as dispersant. Each measurement has been carried out in triplicate and results were averaged and analysed.

2.8.3.3 Differential scanning calorimetry (DSC)

DSC has been performed following the adapted protocols from [48, 82]. Reference and samples have been maintained at same temperatures [83].

Both peptide and protein samples were degassed under vacuum, from 30 minutes to an hour. A Nano DSC (TA Instruments, US) has been used to analyse protein stability. The DSC cells were conditioned twice with buffer for both sample and reference capillaries to

ensure they behaved uniformly. Protein samples were loaded first in the sample capillary and later the buffer was loaded in the reference capillary. Pressure was kept constant at 3 bars, while the temperature scans were carried out from 20 °C to 100 °C at a scanning rate of 1 °C/minute. After the run, the buffer reference heat flow curve is subtracted from the protein sample heat flow curve. Sample preparations were done at concentration gradients of 1 mg/ml, 5 mg/ml to 7 mg/ml for all lengths of peptides. The total enthalpy change for the process is given by the integral under DSC peak, above the baseline. Heat capacities and changes in heat capacities could be determined from the shift in baseline of thermogram.

The software Nano analyze has been used to calculate the melting temperature and enthalpy [84]. The data have been acquired in units of μJ/sec and converted into molar heat capacities in units of $KJ/molK^{-1}$ for further analysis. For enthalpy and melting point calculations, after reference buffer subtractions, the constructed baseline has been added with a sigmoidal baseline to calculate the area under curve. Further, a two state scaled model has been used to fit the raw data with the experimental model.

2.8.4 Transmission electron microscopy (TEM)

2.8.4.1 Grid preparation

The protocols for TEM sample preparation have been followed from Fallas et al and Hartgerink et al.[32, 48]. Uranyl acetate 2 % has been used for staining and stain was required to be stored in dark at 4 °C. Formvar-coated copper grids (200 mesh) (ProSciTech) were used. Cellophane paper and the milli Q water were used [56]. Initially collagen peptides at a concentration of 1 %, 5 %, 10 % and 20 % were prepared to find out the minimum concentration to observe aggregation. Samples were examined at a consistent peptide concentration of 1 mg/ml and 5 mg/ml. Storage temperatures ranged from 4 °C, 20 °C, and 37 °C. Protein and peptide samples were subjected to heating at temperature conditions of 70 °C and 90 °C. Temperatures were stabilised for incubations with temperature blocks and water baths.

2.8.4.2 Procedure for TEM

Grids were placed in the 20 μl peptide sample droplets for 1 minute. These were transferred to 20 μl of water droplet in the next minute. This washing process was repeated two more times. Staining was carried out ; the grids were placed in 2% uranyl acetate stain for 1 minute

and the extra stain was soaked away using cellophane paper. Grids were kept for overnight drying [85].

2.8.4.3 Uranyl acetate stain for TEM

Uranyl acetate was used to negatively stain particles. Nano Van has been obtained as a 2% solution in water at pH 8.0; uranyl acetate was made up as a 1% solution and filtered prior to use through a 0.2 μm syringe filter (EMD Millipore, Massachusetts, USA). Both were stored in the dark at 4 °C to be used within 1 year.

2.8.5 Infrared Spectroscopy (IR spectroscopy):

Infrared spectroscopy has been performed following procedures from [86, 87]. Spectra was recorded in the frequency region of 700 – 4000 cm^{-1}, under resolution of 2 cm^{-1}, scanning speed of 2 mm sec^{-1} and a scanning rate of 64 scans per samples.

2.8.5.1 Instrumentation

The infrared micro spectroscopy (IRM) beam line at the Australian synchrotron was used for the experiments. It combines the high brilliance and high collimation of the synchrotron beam with a Bruker V80v Fourier transform infrared (FTIR) spectrometer and a Hyperion 2000 IR microscope. It gives high signal to noise ratios at a diffracted limited spatial resolution between 3-8 μm. This makes the beam line ideally suitable for the analysis of microscopic samples e.g. small particles and thin coatings on surfaces.

2.8.5.2 Sample preparation

The crystal samples and the self associated samples (nano-scale assemblies) in the unheated and the heated state were analysed by IR spectroscopy. Crystallized peptide samples were prepared in 1:1 ratios of PEG and 10% acetic acid sample solution. 30-40 % PEG 400, 35 % PEG 400, 30 % PEG 550 and 58 % PEG 300 were used. To prepare the solutions of 10 mM PBS and 20 mM PBS, pH 7 was used. The concentrations used have been 7 mg/ml. The heat-treated samples were heated to 70 °C for 8 minutes and cooled down. Once the samples were ready, these were sandwiched between windows, before introducing them into the spectrophotometer. Two kinds of windows were used (1) a 0.5 mm thick CaF_2 window and (2) a diamond window. The diamond window has been used for all samples.

2.8.5.3 IR Analysis

The IR spectra were collected and analysed following previous reports on other self-assembling peptide systems [88]. The spectra were baseline corrected and the water/buffer

signal was removed by subtraction of the pure water/buffer spectrum from the sample spectra. Data analysis has been carried out by OPUS's IR spectroscopy software [89]. Analysis of peptide structural information has been carried out by decomposition of the absorption spectra, either as a sum of Gaussian components or as a sum of spectra assigned to different structures [86]. The entire wavenumber range of 800-4000 cm^{-1} has been analysed [86].

2.8.6 Atomic force microscopy (AFM)

The procedures for AFM on collagen peptides have been followed from Chemoff et al. and modified [90]. Tapping mode AFM has been used in this work.

2.8.6.1 AFM parameters:

Scanning size has been varied within a range of 20 to 100 µm to detect self-assembly and associations. The aspect ratio was kept at 1:1, while the scan angle was altered between 0 to 90 degrees. Scanning was carried out at a consistent sample/lines ratio of 256. Tip velocity was maintained at 20 µm/sec. Sample preparations were done at 0.1 mg/ml, 1 mg/ml and 7 mg/ml to check for triple helices, nanoscale assemblies and aggregates with collagen model peptides in various buffers like water and acetic acid solutions. Glass slides and mica were used for samples while overnight drying was carried out at 4 °C and for room temperature 20 °C [56].

2.9 Determination of protein and peptide concentration

2.9.1 Nanodrop

Protein solutions were analysed using Nano Drop ND-1000 spectrophotometer (Thermo Fischer Scientific, New Zealand). For pure proteins, absorbance coefficient has been determined by using the online software protein calculator (http://christoph-leidig.de/tprot.html) [91]. The calculating operations were based on the number of aromatic residues, tryptophan and tyrosine residues present. It also depends on the presence of disulfide bonds absorbing light at 280 nm [92].

For proteins like type I collagen, absorbance coefficient of 1.0 M^{-1} cm^{-1} has been used. Calibration of the machine was carried out with water and it was blanked with the respective buffer. The absorbance of 2 µl of protein sample was measured. Measurements were carried out in triplicate. With application of Beer's law, nanodrop software calculated the

concentration of proteins automatically. For GPO peptides, the 205 nm UV wavelength was used for analysis.

2.9.2 Serial dilution for peptides

For preparation of accurate and low concentrations of peptides, serial dilutions have been used. For preparation of low peptide concentrations of 0.2 mg/ml, 0.5 mg/ml, stock solutions of 1 mg/ml concentration have been prepared. The basic chemical equation used is given below [93]:

$$C_1V_1 = C_2V_2$$

Equation 2.1

From Equation 2.1, C_1 is the concentration of stock peptide solution, V_1 is the volume to be used up, C_2 is the concentration of peptides to be made up and V_2 is the volume required for it.

2.10 Buffers

2.10.1 Buffers for CD, DSC and DLS

Buffers in Table 2.2, referred from Engel et al., are as follows [48, 94].

Table 2.2 Buffers used for CD, DSC and DLS

Buffer	**Composition**
PBS buffer pH-7.4 [48]	NaH_2PO_4 (0.13 g/100 ml) Na_2HPO_4 (0.14 g/ml) 150 mM NaCl (0.87 g/100 ml)
100 mM PBS buffer pH-7.4	213 mM NaCl (8 g/100 ml) Na_2HPO_4 (1.42 g/100 ml) KH_2PO_4 (0.245 g/100 ml)
High salt 20 mM PBS buffer	213 mM NaCl (16 g/100 ml) 20 mM Na_2HPO_4 (0.288 g/100 ml)

	20 mM KH_2PO_4 (0.048 g/100 ml)
High salt 100 mM PBS buffer	NaCl (8 g/100 ml) Na_2HPO_4 (1.42 g/100 ml) NaH_2PO_4 (0.245 g/100 ml)
10 mM PBS buffer	0.15 M NaCl (M.W= 58.4 g) 10 mM NaH_2PO_4 (MW= 119.8 g) pH was maintained at 7 with HCl and NaOH
10 mM Phosphate buffer	10 mM Na_2HPO_4 (0.163 g/100 ml) 10 mM NaH_2PO_4 (0.116 g/ ml)

Table 2.3 Buffers for varying pH

Buffer	**Composition**
pH – 3.5 Citrate buffer	10 mM Sodium citrate (0.294 g/100ml) (MW=294.10 g/moles)
pH-7 Phosphate buffer	Na_2HPO_4=0.141 g/100 ml $NaH_2PO_4.1H_2O$=0.137 g/100 ml
pH-8 Tris buffer	Tris=0.121 g/100 ml

From Table 2.3, buffers of various pHs have been selected according to the pK_a values of the peptide terminals.

2.10.2 Mass Spectrometry buffers

Table 2.4 Buffers for Mass Spectrometry samples

Tris buffer	Tris – 0.315 g/100ml

2.10.3 Crystallography solvents

Table 2.5 Solvents for crystallography (v/v %)

PEG	Percentages
PEG 200	10 – 100 %
PEG 300	10 – 100 %
PEG 400	10 – 100 %
PEG 550	10 – 100 %
PEG 600	10 – 100 %
PEG 1000	10 – 50 %
PEG 1500	10 – 50 %
PEG 2000	10 – 50 %
Acetic acid	10 %

For crystallography, the percentages details shown in Table 2.5 have been prepared as follows:

For preparation of 100 µl of 10 % PEG solution, 10 µl of PEG was added to 90 µl of water. For preparation of 20 % of the 100 µl of PEG solution, 20 µl of PEG was added to 80 µl of water etc.

2.11 Crystal formation

2.11.1 Practices and conditions for crystal growth

2.11.1.1 Seeding

Seeding is a technique for separating nucleation and growth. Previously nucleated crystals have been used as seed crystals. These are placed in new drops, equilibrated at super saturation [95]. Two kinds of seeding have been used in this project work (a) micro seeding and (b) macro seeding.

2.11.1.1.1 Micro seeding

The already crystallized peptides along with their droplet solutions were picked up with pipettes. These were the seed droplets. These seed droplets were mixed with 50 µl of mother liquor. The mother liquor was a solution comprising of PEG and 10 % acetic acid. The mixture is vortexed. The dropping solutions have been prepared as shown in Figure 2.1. From the vortexed solution, 5 µl of it was added to fresh 45 µl mother liquor. From this mixture, 5 µl was taken out and added to another fresh Eppendorf with 45 µl of mother liquor in it. Again 5 µl was pipetted out and added to new 45 µl of mother liquor solution. This has been repeated till the desired numbers of dilutions were prepared [95].

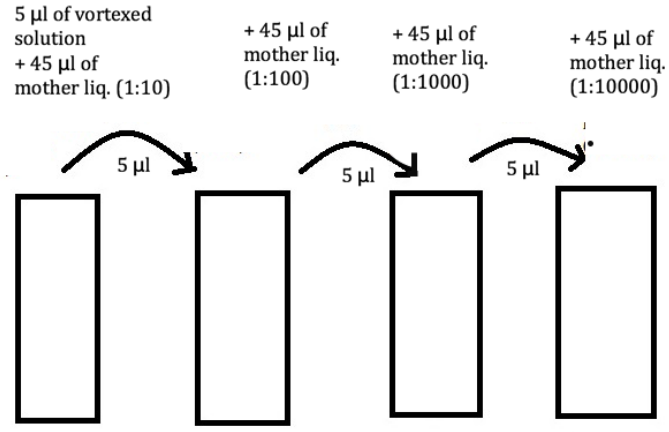


Figure 2.1 Micro seeding, solutions of ratio 1:10, 1:100, 1000 and 1:10000 were prepared

From each Eppendorf's mixture, 0.6 µl of the solution was pipetted out and was placed on a cover slip. As shown in Figure 2.2, along with 0.6 µl of each diluted seeding solution, 1.5 µl of the peptide solution and 1.5 µl of the reservoir solution formed a hanging droplet.

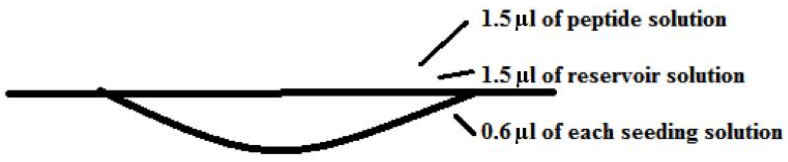


Figure 2.2 Micro seeding with hanging drop vapor diffusion method. Droplet comprises of peptide solution, reservoir solution and seeding solution

2.11.1.1.2 Macro seeding

From crystallized peptides in the droplet, 2 µl of droplet was placed in the 50 µl of the reservoir solution. It was vortexed for 2 minutes. The above solution was dissolved with 200 µl of the reservoir solution and vortexed again for another 2-5 seconds. This is the seed solution [95].

Peptide solutions of 2 µl, 2 µl of the reservoir solution and 0.2 µl of the prepared seed solution were placed as the droplet on the cover slip as shown in Figure 2.3.

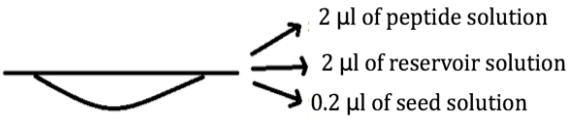


Figure 2.3 Hanging drop, macro seeding; droplet comprises of, peptide solution, reservoir solution and seed solution

2.11.2 Optimisation of crystallization of collagen mimetic peptides

For determination of ideal conditions for crystal formations for collagen model peptides, factors for screening were selected. Those were (1) temperature, (2) time, (3) PEGs, (4) PEG diluted percentage solutions, (5) peptide concentrations, (6) pH and (7) chain lengths.

2.11.2.1 pH screening

The peptides were weighted and dissolved in 10 % acetic acid solution. 10 % acetic acid (v/v) was prepared by adding 1 part (ml) of acid and 9 parts (ml) of MilliQ water.

Reservoir volume wells of 24 well hanging drop plate were PEG based buffers. The selected pH were pH 2 to pH 4.5. Peptides in acetic acid buffers, which were optimized at various pH were mixed with the reservoir contents in equal ratios. This mixture was placed as a hanging droplet on cover slips and each of the wells were covered. The crystal trays were incubated at their respective temperature conditions. Keeping the rest of the factors like time, temperature, reservoir buffer, peptide concentration, and chain length constant, the crystal plates were checked for crystal formations every day.

2.11.2.2 Temperature screening

After setting up crystal droplets, the hanging drop and sitting drop plates were incubated at mainly three different temperatures, at 4 °C, room temperature of 20 °C and at 37 °C. Crystals appeared fastest at 4 °C and at room temperature conditions, compared to 37 °C.

2.11.2.3 Peptide concentration screening

Peptide concentration screens of all (GPO) peptides, from 1 mg/ml to 15 mg/ml were prepared as shown in the Table 1 of peptides. Crystals have been reported for $(GPO)_5$ peptides at 15 mg/ml with 35% PEG 400. Crystals appeared for $(GPO)_7$ peptides at a concentration of 15 mg/ml with 35 % PEG 400. $(GPO)_{10}$ peptides reported crystals at 5 mg/ml with 22 % PEG 400.

For $(GOO)_{10}$ peptides, at a concentration of 7 mg/ml and 22 % PEG 400 composition, crystals were formed. The base solvent chosen to dissolve the peptides was 10 % acetic acid solution at pH 2.3 and this was mixed with equal portions of reservoir PEG solutions. Droplets were placed on the cover slips and inverted cover slips were placed on the wells of the hanging drop plate.

2.11.2.4 Mother liquor

Peptides were dissolved in various buffers i.e. 10 % acetic acid, PBS buffer or phosphate buffer. Once the crystals were formed, the buffer yielding crystals were used for further screenings.

2.11.2.5 PEG screen

Optimization of crystals were performed with various PEG solutions, such as PEG 200, PEG 300, PEG 400, PEG 550, PEG 600, PEG 1000, PEG 1500 and PEG 2000 for the selected GPO peptides. Every single PEG has been diluted to ten different % solutions, starting from 10 % up-till 100 %. For example, 10 % PEG 200, 20 % PEG 200 till 100 % PEG 200, for the 10 wells in a crystal plate. Same preparation was repeated for rest of the PEGs. All these were prepared in MilliQ water. The various PEG plates at all percentages were checked for crystal formations every day, (for around 3 to 4 weeks). The unsuccessful wells would show precipitate formations. After the nearest crystal formations were recognized in-between two consecutive percentage wells, for example 20 % PEG 400 and 30 % PEG 400, further optimizations were performed. The % window is expanded to another 5 wells to obtain the best diffraction quality crystals. Hence wells of 20 %, 22 %, 24%, 26%, 28 % and 30 % of

PEG 400 would be prepared and incubated for the next couple of weeks. These were checked every day. The best percentage dilutions of PEG yielding crystals, have been taken into consideration for consistently making diffraction quality crystals.

2.11.2.6 Lengths of peptide

By varying chain length of peptides, optimization of crystal formations were carried out.

2.11.3 Conditions for crystallization

All crystals have been obtained in 10 % acetic acid buffer at pH 2.4.

2.11.3.1 $(GOO)_{10}$

$(GOO)_{10}$ formed crystals at 7 mg/ml peptide concentration in 10 % acetic acid solution in 22 % PEG 400 at incubations of 20 °C and also at 4 °C. Heated crystals used for IR spectroscopy were obtained at the same conditions.

2.11.3.2 $(GPO)_{10}$

$(GPO)_{10}$ crystals formed at 5 mg/ml, 7 mg/ml and 15 mg/ml concentrations at 4 °C with 22 % PEG 400. The crystals were also obtained at a concentration of 4 mg/ml concentrations at 4 °C within two weeks of incubation time.

2.11.3.3 $(GPO)_7$

These formed crystals at 15 mg/ml concentration, at 30 % and 40 % PEG 400 at room temperature, 20 °C and also at 4 °C.

2.11.3.4 NH_2-$(GPO)_7$-NH_2

These formed crystals at PEG conditions of both 30 % PEG 550 and at 58 % PEG 300. These were achieved at 4 °C, at both concentrations of 10 mg/ml and 15 mg/ml.

2.11.3.5 $(GPO)_5$

These crystalline precipitates were formed at concentrations of 15 mg/ml and 20 mg/ml with 35 % PEG 400, at an incubation time of 1 week at 4 °C.

2.11.3.6 $(GPO)_3$

These precipitates formed at 15 mg/ml mostly and also at 20 mg/ml concentrations with PEG 400 at percentages of 22 %, 25%, 30 % and 35 %. These formed 4 °C and also at room temperature of 20 °C.

2.11.3.7 Racemic $(GPO)_5$-Gly

These crystals formed at 15 mg/ml and also at 7 mg/ml concentrations. 58 % PEG 300 at either of the incubation temperature of 4 °C or 20 °C was required, along with the incubation times of 2 to 3 weeks.

2.11.4 Washing crystals for mass spectrometric analysis [96]

Preparation required several crystal droplets (usually four drops). For mother liquor, 10 % acetic acid and PEG solutions were used. 50 µl of the mother liquor was kept aside in a 1.5 ml Eppendorf tube. Crystals were transferred to an Eppendorf tube by a pipette. Contents of the Eppendorf tube were centrifuged at a maximum speed of 18,000 rpm at 4 °C. The supernatant was discarded. To this base pellet, 100 µl of mother liquor was added. It was vortexed and centrifuged for 10 minutes at a maximum speed of 18000 rpm at 4 °C. The supernatant was discarded and this process was repeated from five to seven times. The crystals were dissolved in the preferred solutions.

2.11.5 Mosquito crystallization robot

Mosquito crystallization is an ideal liquid handling tool for setting up protein/peptide crystallization experiments. It places droplets very precisely. The sitting drop method on the 96 well sitting drop plates used around 300 nl to 500 nl peptide droplets in each well, while the reservoir carried 40-50 µl of reservoir liquid.

2.12 X-ray Diffraction

The diffraction data were obtained for $(GPO)_{10}$ and racemic $(GPO)_5$-Gly crystals. Crystals were taken to the Australian Synchrotron facility for data collections. The scattered X-rays form a characteristic pattern containing information about the positions of atoms in the structure. The scattering pattern is the Fourier transformation of the electron density in the molecule [97].

In this work, both the MX1 and MX2 beam-lines have been used. The MX1 beam line was most efficient for peptide crystals. For the data collection on (GPO_5)-Gly and $(GPO)_{10}$ crystals, the energy range was fixed between 7 to 18 keV. The wavelength used for MX1 was 0.9537 Å. The wavelength was kept within the range of 1.77 to 0.69 Å. The un-collimated beam was of size 120 µm × 120 µm.

2.12.1 Data processing for $(GPO)_{10}$

1440 images were collected. Images were indexed, integrated and scaled with XDS [98]. The resolution range was determined by considering the completeness (over 90 % and 95 % in overall and in the highest shell respectively), 0.8 Å was the selected resolution for $(GPO)_{10.}$ Data was scaled in SCALA or AIMLESS in CCP4. The % of reflections selected as R free set was 0.05 from the dataset. File conversion from the mtz format of CCP4 to the SHELX format, was carried out in CCP4. Structure was determined by direct methods using SHELX [99]. SHELXD was used. All models were refined with REFMAC5 [100]. The model was refined anisotropically. Addition of water, was seen to decrease the R factor and R free, very significantly. Manual adjustments of the structure were performed with COOT [101]. The torsion angles phi (φ, indicative of C'-C' distance), psi (ψ, indicative of N-N distance), and omega (ω, indicative of Cα-Cα distance) were calculated using DSSP [102] and *PROCHECK* from *CCP4* suite [103]. All of the structural figures were generated with PyMOL (DeLano Scientific LLC).

2.12.2 Data processing for $(GPO)_5$-Gly

For $(GPO)_5$-Gly racemic crystals, three procedures were attempted to solve the structure

2.12.2.1 Direct method (using SHELXD) and small molecular crystallography

Data processing was initially carried out with XDSME. BRUTE command was used for a better integration of the dataset. From XDSME, the .INP file was generated and changes were made into the .INP file to include the selected probable space group and the selected probable cell parameters. Selection was made from the XDS suggested space groups and cell parameters, in the CORRECT file. XDS was run with selected space group and cell parameters. The .HKL file was further scaled with XSCALE. The XSCALE generated XSCALE.HKL file was converted to a SHELX.HKL file. SHELXD was run with the .HKL file to generate structural solutions in forms of .pdb, .res and log based .ins files.

These were used for further refinements with Olex2. A clear solution could not be obtained with Olex2 after small molecular method based manual refinements.

2.12.2.2 Molecular replacement

After XDS processing, the .mtz file of raw reflections was obtained by a conversion of the .HKL file in CCP4's COMBAT. CCP4's imosfilm was used to index and integrate the dataset. AIMLESS or SCALA was used to scale the dataset. CCP4's Molrep was used to find a solution with Molecular Replacement. Collagen model peptide models like 1V4F, 1V7H

from PDB, and the solved structure of $(GPO)_{10}$ from this research work, were used as potential models. Refinements were carried out with CCP4's REFMAC5.

2.12.2.3 Direct method (using SHELXD) and model building

Processing was initially carried out with XDSME to generate an .INP file. From the XDS generated probable space groups and cell parameters in the CORRECT file, the most probable space group and parameters were selected and updated in the XDS.INP file. XDS was run for data processing, indexing and integration. It was followed by XSCALE for scaling. The generated XSCALE.HKL file was converted to SHELX.HKL. SHELXD was run to find a solution. The generated solution from SHELXD was further refined in REFMAC5 along with the XSCALE generated reflections file. Models of space group C2 and I4 showed a triple helical skeleton. But a detailed atomic structural model could not be obtained from SHELXD to clearly identify the proline and hydroxyproline rings. This problem could have occurred at the diffraction data collection. Probably the diffraction data set was not complete (> 90% collection), because of the crystals being radiation sensitive.

3 Effect of varying chain lengths of GPO peptides on self-assemblies

In this chapter the impact of varying chain lengths of GPO peptides (glycine proline hydroxyproline peptides) on triple helix formations and triple helical stabilities have been studied. The results of in-depth characterizations and examinations on the triple helical assembly and the supramolecular assembly formations are discussed.

3.1 Self Assembly

Self-assembly is defined as the spontaneous organization of molecular building blocks, which forms well-defined and stable macroscopic structures stabilized by non-covalent bonds (H-bonds, water mediated H-bonds, ionic bonds, hydrophobic and van der Waals interactions) [104]. Short peptides are easy to design and synthesize, making them an excellent model system for studying self-assembly.

Several researches have focused on the self-assembly of collagen mimetic peptides[105]. These assembled peptides have found uses as biomaterials that have proved to be of considerable importance for applications like scaffolding for tissue repair in regenerative medicine, drug delivery, biological surface engineering etc [104, 106, 107].

Deriving inspiration from the aforementioned literature, GPO peptides [$(GPO)_3$, $(GPO)_5$, $(GPO)_7$, $(GPO)_{10}$ peptides] were picked for experimentation and the impact of varying chain lengths on triple helix formations, triple helical stabilities and supramolecular assembly formations is studied.

3.2 Characterization studies on triple helical self-assemblies and supra-molecular assemblies from literature

3.2.1 Triple helical self-assemblies

Collagen triple helical stability is governed by parameters including hydration, chain length, hydrogen bonding and sequence [108]. From literature, it is seen that the first order assemblies, triple helices of collagen mimetic peptides and the melting points of these triple helices have been investigated via circular dichroism and differential scanning calorimetry [109].

Differential scanning calorimetry was used to understand the thermodynamics and kinetics of triple helical assembly with measurements of transition temperature, enthalpy and heat capacity change [109]. Dynamic light scattering has been used to study the self-association dynamics of peptides along with their homogeneity and aggregation. The analysis of the triples helical assemblies, supramolecular assemblies and fibrillar height measurements has been carried out with atomic force microscopy. Transmission electron microscopy has been used to visualise fibrillar associations.

Triple helical domains are present in all supramolecular assemblies [48]. Several reports on collagen like peptides that self-associate into supramolecular assemblies e.g $(POG)_{10}$ peptides have been studied [110]. Formation of liquid crystalline arrays by the triple helical peptides have also been studied [48, 111].

3.2.2 Peptide selection

3.2.2.1 Literature search for peptide selection

All homotrimer motifs until date, discussed in the introductory chapter of the thesis are based on the triad- proline-hydroxyproline-glycine (POG) as the repeating unit, these sequences can be further subdivided into four main types of designs -

i) Simple trimeric motifs;
ii) C-terminal appended trimeric motifs;
iii) Sequences with N-and C-terminal trimeric flanking motifs;
iv) Sequences with N-and C- terminal "sticky ends" trimeric flanking motifs.

The first type of design is based on ten repeats of a single triad. The POG triad was found to assemble into triple helices in solution or when crystallized, in a similar way to the PPG motif. Two POG variants have also been reported to assemble into triple-helices in solution, namely the OPG triad and a fluorescent variant of POG [48, 112, 113].

Holmgren et al. have performed experiments to study the abnormalities of fibrillar arrangements of collagen, which can give rise to serious diseases [113]. Similarly, Kar et al. have performed experiments with collagen peptide systems and characterized them, which could lead to understanding the physiological self-association of collagen molecules [48].

In-depth characterization of the chain lengths of collagen mimetic GPO peptide systems (simple trimeric motifs) that influence the propensity to form triple helical and higher order assembly structures have been discussed in this thesis. This research has been performed because literature suggests that there have been many studies on peptides of $(POG)_{10}$, $(PPG)_{10}$, $(GOO)_{10}$ etc [69, 70, 109, 114], but there have been no studies on the effect of varying GPO chain lengths on self-assembly.

The second type of design incorporates two to three residues, mainly aromatic and/or hydrophobic residues, at the N and C-terminal of sequences made of 9 to 10 repeats of the POG/GPO triad (sequences 5-10). At this specific position, aromatic residues have been found to promote higher order assembly of the POG triple-helices into microfibrils in solution [56]. It has also been observed that, successful formation of micro fibrils by POG peptides are favored by aromatic residues incorporated at the N and C termini.

The third type of design is based on sequences that are 30 residues long, that incorporate 3-4 repeating triad sequences of (GPO) or (POG) on both sides, N-terminal and C-terminal, of a small 3 or 9 residue long sequence. The presence of basic repeating units from collagen, that joins proline, hydroxyproline and glycine residues as triads on both ends, enhances the triple helix formation. However, these helices do not form higher assemblies, as supported by previous work done on the T3-785 motif ($(POG)_3$ITGARGLAGPOG$(POG)_3$) with low imino acid content. When incorporated with two flanking regions of imino acid rich motifs on either sides, the CMP exhibits self assembly [46, 115].

The fourth type of design which is based on sequences that are 36 residues long, comprise of three triads each repeating four times. $(PRG)_4(POG)_4(EOG)_4$, was designed with a sticky-end motif, so that at pH 7, the arginine containing N-terminal region would form electrostatic interactions with the glutamic acid containing C-terminal region, which forms a triple helix and subsequently microfibrils [32, 43].

3.2.2.2 Peptides selected for the thesis

The presence of proline and hydroxyproline stabilises the individual polyproline II like helices in collagen triple helices. The triple helical chains have a repeating structure of Gly-Xaa-Yaa, in which Xaa and Yaa are proline and hydroxyproline respectively. Glycine is present in every third residue along the chain that facilitates winding in collagen [31].

Deriving inspiration from literature, $(GPO)_n$ peptides have been picked for this thesis work, where n is 3, 5, 7 and 10. These n values were selected because aim of this piece of research work was to crystallize and solve the shortest possible GPO collagen. An amidated or blocked sequence of $(GPO)_7$, denoted as NH_2- $(GPO)_7$-NH_2 has also been synthesized to examine the effect of amidation on the $(GPO)_7$ peptides. All peptides used in the experiments have been HPLC treated by Mimotopes, maintaining a purity of above ~ 95 %. From previous research, changes brought about in the helical self assembling properties of $(GPO)_7$ peptides due to amidation have not been researched before. The amidated peptides, NH_2-$(GPO)_7$-NH_2 have been easier to synthesize than the NH_2-$(GPO)_7$-COOH peptides, by the Maurice Wilkins Centre (Auckland) and thus the assembly differences for both peptides have been studied.

3.3 Results and Discussion

Circular dichroism (CD) is the branch of absorption spectroscopy that is used to measure the interaction of polarised light with asymmetric molecules [116]. CD spectra of proteins and peptides are usually measured in two spectral regions - near UV and far UV. Bands in the far UV region (below 250 nm) represent the primary electronic transitions of the amide groups in proteins. The signs, magnitudes and positions of these bands are strongly dependent on the θ and Ψ angles of the peptide bond.

A plot of ellipticity vs wavelength gives information about the secondary structure of the protein. A positive peak (maxima) at 225 nm depicts the presence of triple helices, while a negative dip (minima) at around 200 nm shows the presence of random coils [117] [80]. Triple helix is also depicted by the dominant maximum at around 225 - 235 nm.

Beer Lambert law states that the extinction coefficients, εL (ƛ) is related to the ordinary absorption εLR by the formula in Equation 3.1

$$\varepsilon LR = ALR(ƛ)/lc$$

Equation 3.1

Where, l is the path length of the cell (in cm) and c is the concentration in moles/L.
Hence, the Equation 3.2 below can explain circular dichroism

$$\Delta\varepsilon(\lambda) = ALR(\lambda)/lc$$

Equation 3.2

$$ALR(\lambda) = AL(\lambda) - AR(\lambda) \qquad [118]$$

Equation 3.3

CD is not usually expressed as $\Delta\ \varepsilon\ (\lambda)$, but as molar ellipticity $\theta\ (\lambda)$, which has units in deg cm^2 decimol $^{-1}$. Molar ellipticity is related to $\Delta\ \varepsilon\ (\lambda)$ by the formula in Equation 3.4

$$\theta(\lambda) = 3298\Delta\varepsilon(\lambda)$$

[118]

Equation 3.4

Generally, the intensity of the CD signal increases with the optical density (concentration of sample). The profiles for the helix, sheet and turn elements of the secondary structure have been referred from (Besley and Hirst 1999). Mean residue ellipticity (MRE) has been calculated by the formula in Equation 3.5

$$\text{Mean residual ellipticity } [\theta] = \frac{\text{MRW} \times \theta_{\text{obs}}}{(10 \times \text{d} \times \text{c})}.\text{deg}.\text{cm}^2.\text{dmol}^{-1}$$

[119]

Equation 3.5

Where, θ obs = observed ellipticity in degrees

d = path length in cm

c = concentration in g/ml

And Mean residual weight (MRW) has been calculated from molecular mass/ (N – 1),

N = number of amino acids, while for most proteins MRW is taken to be around 110 (Kelly et al. 2005).

To calculate MRE, MRW was assumed to be 110 for all GPO collagen mimetic peptides of varying chain lengths in this thesis work.

3.3.1 Critical concentration determination for triple helix formation

For varying chain lengths of GPO peptides i.e. $(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ peptides, the maximum concentration required for helix formations by CD is 1 mg/ml. [48] report that $(POG)_{10}$ peptides propel self-associations and bring about aggregations, above 3 mg/ml concentrations with incubation times of 4 hours. The same was observed with the studies on the (GPO) peptides in this chapter. Disturbances were observed beyond concentrations of 1 mg/ml. Incubation time period has been chosen to be comparatively longer (3 days instead of 4 hours) for these experiments to give enough time for self-assembly to fully take place.

Therefore, concentrations of 0.2 mg/ml, 0.5 mg/ml and 1 mg/ml were picked to conduct the experiments. The concentrations were prepared in 20 mM PBS buffer (150 mM NaCl salt) and incubated for 3 days at 4 °C at pH 7. This protocol was chosen and modified from the literature [48, 109]. Concentration units of milligrams per milliliters have been used instead of molar concentrations (moles per litres) for the peptide solutions because milligrams per milliliters is the measure of density of the solution. From studies of [120], concentration units were used in terms of milligrams per milliliters to examine and compare stabilities of the triple helical (POG)n and (PPG)n collagen mimetic peptides.

Concentrations of 0.2 mg/ml of (GPO) peptides show the least MREs in Figure 3.1A. The 0.5 mg/ml concentrations of peptides show higher triple helical MREs than the samples at 0.2 mg/ml concentrations (see Figure 3.1B). The highest MRE signals are recorded at 1 mg/ml (Figure 3.1C).

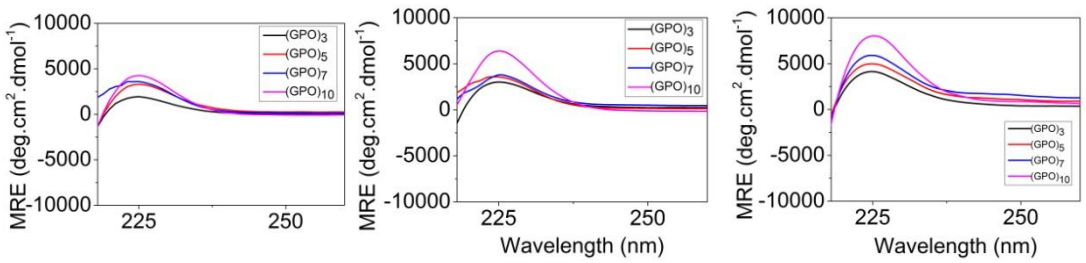


Figure 3.1 CD analysis; Critical concentration determination for triple helix formations of $(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ peptides, incubated for 3 days at 4 °C at pH 7 in 20 mM PBS buffer (150 mM NaCl salt) (A) 0.2 mg/ml (B) 0.5 mg/ml (C) 1 mg/ml.

The triple helix formation was found to be directly proportional to the concentration of peptides. The MREs are indicative of triple helical content and were seen to increase with the increasing chain lengths of peptides for all GPO peptides. Since 1 mg/ml had the most helical content for all the peptides, this concentration was used for further experiments.

3.3.2 Determination of most effective pH and triple helix formations

From previous studies, pH 7 has been seen to be highly supportive of triple helix formations for collagen mimetic peptides like $(GPO)_{10}$ and $(GPO)_8$ [48, 121].

Solutions of $(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ in citrate buffer (pH 3.5), 20 mM PBS buffer (pH 7) and tris buffer (pH 8) each were prepared. These were prepared at 1 mg/ml concentrations and were incubated at 4 °C for time periods of 3 days. The buffers had been previously filtered to remove impurities.

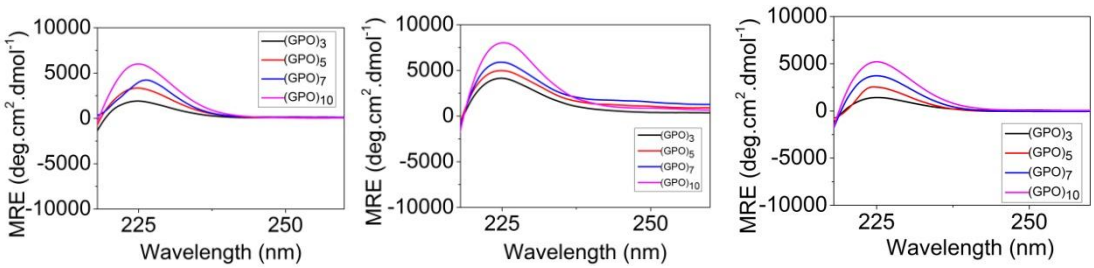


Figure 3.2 CD analysis; Effect of pH on triple helix formation for $(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ peptides, incubated for 3 days at 4 °C. (A) pH 3.5 in citrate buffer (B) pH 7 in 20 mM PBS buffer (C) pH 8, Tris buffer

It was observed that triple helical content was maximum at pH 7 for most of the peptides. $(GPO)_{10}$ peptides show similar CD signals for all the three pH 3.5, 7 and 8. $(GPO)_7$, $(GPO)_5$ and $(GPO)_3$ peptides exhibit maximum ellipticity at pH 7 (Figure 3.2A).

Helix formations were least visible at pH 8 (Figure 3.2C). Based on these results, further experiments were conducted at pH 7.

3.3.3 Effect of time on triple helix formation

Solutions for $(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ peptides were prepared in 20 mM PBS buffer (150 mM PBS) each at pH 7 at 1 mg/ml. These were incubated at 4 °C for the time periods of 1 hour, 4 hours and 3 days respectively. The buffer was filtered to remove unwanted particles. These time periods were picked by following protocols from [48]. At incubation time period of 3 days (Figure 3.3C), maximum ellipticity was observed for all

peptides; it was followed by samples incubated for a time period of 4 hours (Figure 3.3B). Samples at 1 hour of incubation time (Figure 3.3A) show the least ellipticities.

Helix formation was found to be directly proportional to time. Further measurements were therefore carried out after 2 days of incubation, after which equilibrium had been reached.

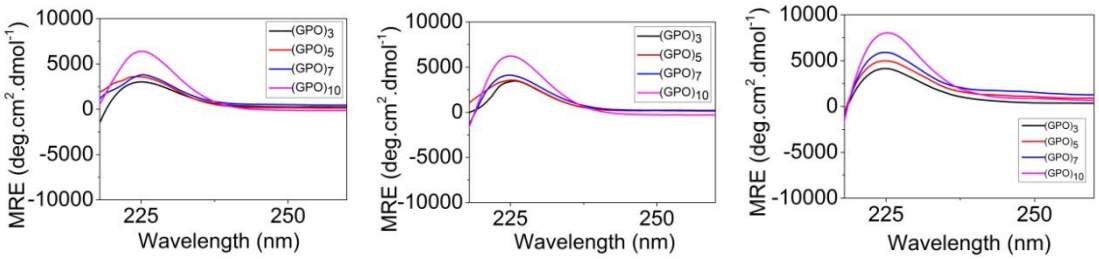


Figure 3.3 CD analysis; Effect of incubation times on triple helix formation for $(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ peptides, prepared in 20 mM PBS buffer (150 mM NaCl salt) incubated at 4 °C, at pH 7 (A) 1 hour (B) 4 hours (C) 3 days

3.3.4 Triple helical stabilities - Analysis by CD

Among the various collagen model peptides of (Gly-Xaa-Yaa)n make up, which were prepared, the (GPO) triad was noticed to stabilize triple helix formations the most (where n= number of triplet repeats). The melting point (MP) of $(GPP)_{10}$ peptides was reported to be 28 °C and the MP of $(GPO)_{10}$ was ~ 60 °C [60, 122]. For comparison, the MP of various collagen peptides reported in literature is tabulated in Table 3.1.

Table 3.1 CD melting temperatures of various collagen peptide helices from literature have been reported [48]

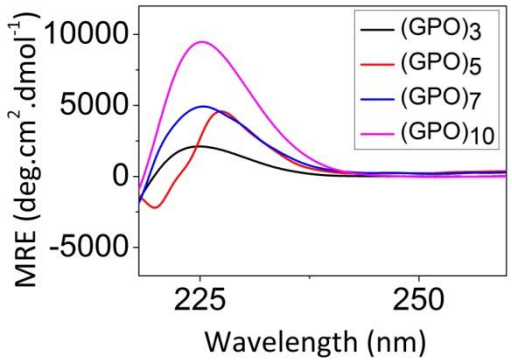
(GPO)n (Literature)	Melting Temperatures of helices
$(POG)_7$	25 °C
$(POG)_8$	47 °C
$(POG)_{10}$	60 °C
$(POG)_{12}$	70 °C

In this section, effect of varying chain lengths of GPO peptides on melting of triple helices has been discussed. Both automated and manual melting experiments have been carried out. In automated melting, the CD instrument and the software automatically control the temperature increments within a provided range of temperature, throughout the process of melting. From the data of previous experiments, 225 nm was picked for all temperature scans in the melting analysis of triple helices in this research. The instrument automatically performed the temperature scans at regular time intervals, at a provided scanning rate of 1 °C/minute, for both heating and cooling processes. In manual melting, instead of the temperature controller and the CD software taking automated readings, manual wavelength scans were performed alongside the manually controlled temperature increments/decrements, within the manually controlled time intervals at scanning rates of 1 °C/minute.

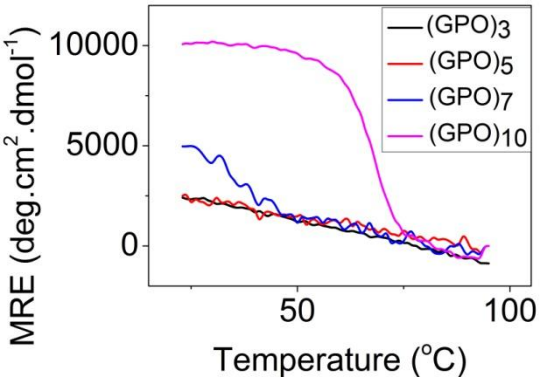
3.3.4.1 Automated melting analysis of triple helices

Peptide sample solutions of $(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ were prepared at 1 mg/ml at room temperature (20 °C) and were incubated for 2 days in water. Thermal unfolding was carried out according to the following 5 steps explained below (see Figure 3.4).

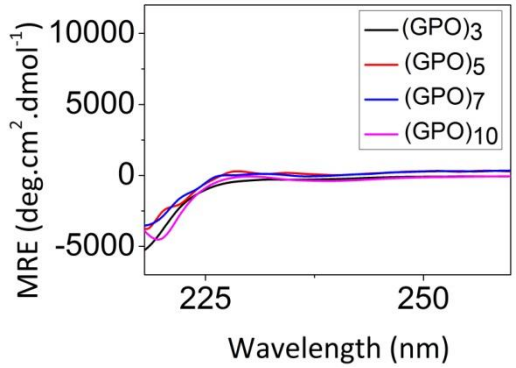
In the first step triple helix formation was characterized at 20 °C. Plots of MRE vs wavelength were measured. These wavelength scans for peptides were carried out from 260 nm to 180 nm at 1 °C/minute (Figure 3.4A). The helix formations were recorded to be highest for $(GPO)_{10}$ peptides, followed by $(GPO)_7$, $(GPO)_5$ and least for $(GPO)_3$ peptides. The propensity of helix formation was found to be directly proportional to chain length.



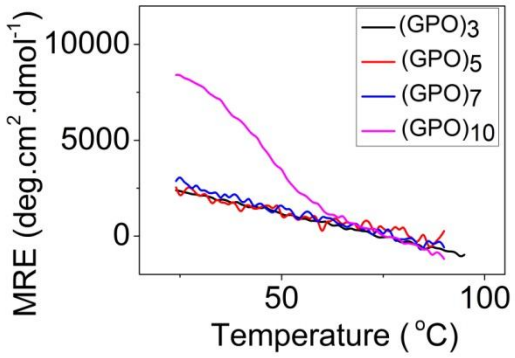
A



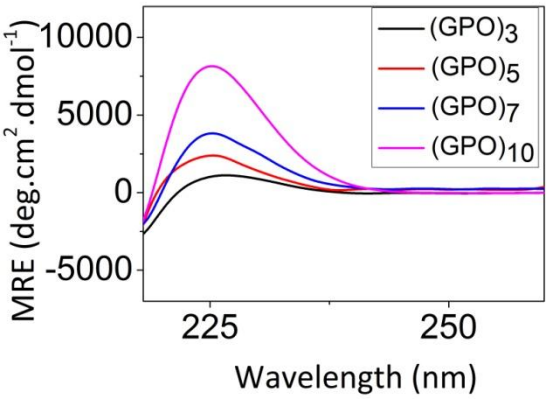
B



C



D



E

Figure 3.4 CD analysis; Automated melting analysis of triple helical unwinding and rewinding for $(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ peptides in water for 2 days, at room temperature (20 °C). (A) Wavelength scan at 20 °C (B) Temperature scan from 20°C to 90 °C, at 225 nm (C) Wavelength scan at 90 °C (D) Temperature scan from 90°C to 20 °C at 225 nm (E) Wavelength scan back at 20 °C

In the second step, temperature scans were carried out (Figure 3.4B). Temperature was gradually increased at a scanning rate of 1 °C/minute. Changes in MRE were recorded at 225 nm in the temperature range from 20 °C to 90 °C. Triple helical melting was observed for all peptides. For $(GPO)_{10}$ peptides, the melting temperature range was seen at around ~ 65 °C. Buhot et al define the cooperative transition curve as the temperature range within which the major transitions occur and is also the range of melting for triple helices [32, 56, 123]. The cooperative transition curve for melting is shown in pink for $(GPO)_{10}$ peptides in Figure 3.4B. No cooperative transition curves were observed for rest of the peptides, but the process of melting was clearly visible.

In the third step, MRE vs wavelength scans were performed at 90 °C to check for the presence of triple helices. The triple helices had disappeared for all peptides (see Figure 3.4C). It is believed to have happened due to the helices unfolding and changing into monomers at 90 °C.

In step four, a reverse temperature scan was performed. The temperature was gradually decreased from 90 °C to 20 °C at a scan rate of 1 °C/minute. In this region, changes in ellipticities were observed (see Figure 3.4D) and helical rewinding was observed for all peptides.

In the fifth and final step, a further scan of MRE vs wavelength was performed at 225 nm at 20 °C (see Figure 3.4E). The triple helices for all GPO peptides re-appeared again. This points in the direction of reversibility in conformation for all the selected GPO peptide systems.

The MPs observed in this section have been tabulated in the Table 3.2. Melting range of $(GPO)_{10}$ peptides was around ~ 65 °C from CD.

3.3.4.2 Manual melting analysis of triple helices

Peptide solutions of $(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ were prepared in 20 mM PBS solution at pH 7 at 4 °C incubation for 2 days.

Table 3.2 Stability studies of (GPO)n; melting temperatures from this chapter, have been established here (where n is the number of triplet repeats)

	Selected peptide systems for thesis work	Melting temperatures (°C)
Automated CD melting studies in water		
	$(GPO)_3$	-
	$(GPO)_5$	-
	$(GPO)_7$	-
	$(GPO)_{10}$	~ 65 °C
Manual CD melting studies in 20 mM PBS (150 mM NaCl) pH 7		
	$(GPO)_3$	-
	$(GPO)_5$	-
	$(GPO)_7$	35- 40 °C
	$(GPO)_{10}$	65-70 °C
Automated CD melting studies in 100 mM PBS (213) mM NaCl) pH 7		
	$(GPO)_3$	-
	$(GPO)_5$	-
	$(GPO)_7$	35-40 °C
	$(GPO)_{10}$	65-70 °C
DSC melting studies In 20 mM PBS, (150 mM NaCl) pH 7		
	$(GPO)_3$	-
	$(GPO)_5$	-
	$(GPO)_7$	33.8 °C
	$(GPO)_{10}$	~ 63 °C

Apart from performing the melting analysis of GPO peptides in water (in the above section), melting was also analysed with 20 mM PBS buffer. Automated melting analysis was carried out for $(GPO)_3$ and $(GPO)_5$ peptides. Because of an unexpected break down of the automated temperature regulator controller, manual-melting analysis experiments were performed for the rest of the peptides (Although these did not alter the results).

For $(GPO)_3$ peptides, the triple helix formations were observed at room temperature i.e. 20 °C. At 90 °C, it was observed that the helices disappeared due to melting. Triple helices reappeared again, after cooling back to 20 °C (Figure 3.5A). If the analysis of $(GPO)_3$ peptides here seem different from the study on $(GPO)_3$ peptides in the above section (Figure 3.4A), it would be because the readings above were taken in water, whereas the study of manual melting in this section are carried out in PBS buffer (consisting of 150 mM NaCl salts).

For $(GPO)_5$ peptides, triple helix formation was observed at 20 °C. The helices disappear at 90 °C. It was interesting to note that an unknown feature/structure formed as shown by the dip at 225 nm (Figure 3.5B). No reports have been seen on this feature/structure in literature. On cooling back to room temperature i.e. 20 °C, the helices reappeared again.

Manual melting analysis on $(GPO)_7$ and $(GPO)_{10}$ peptides have been performed. The disappearance of the maxima at 225 nm symbolizes the conversion of triple helices to random coils. With cooling, the re-appearance of the maximum at 225 nm was the point of re-winding of the random coiled monomers to triple helices; random coils are usually depicted by minimum near 198 to 200 nm [124].

It was observed that during the melting of the $(GPO)_7$ and $(GPO)_{10}$ peptides and the eventual disappearance of the maximum at 225 nm, the minimum also shifted to lower wavelengths. During these changes a point was recorded which symbolizes two-state equilibrium between triple helical state and the random coil state. This specific wavelength is known as the ‘isodichroic point’[60].

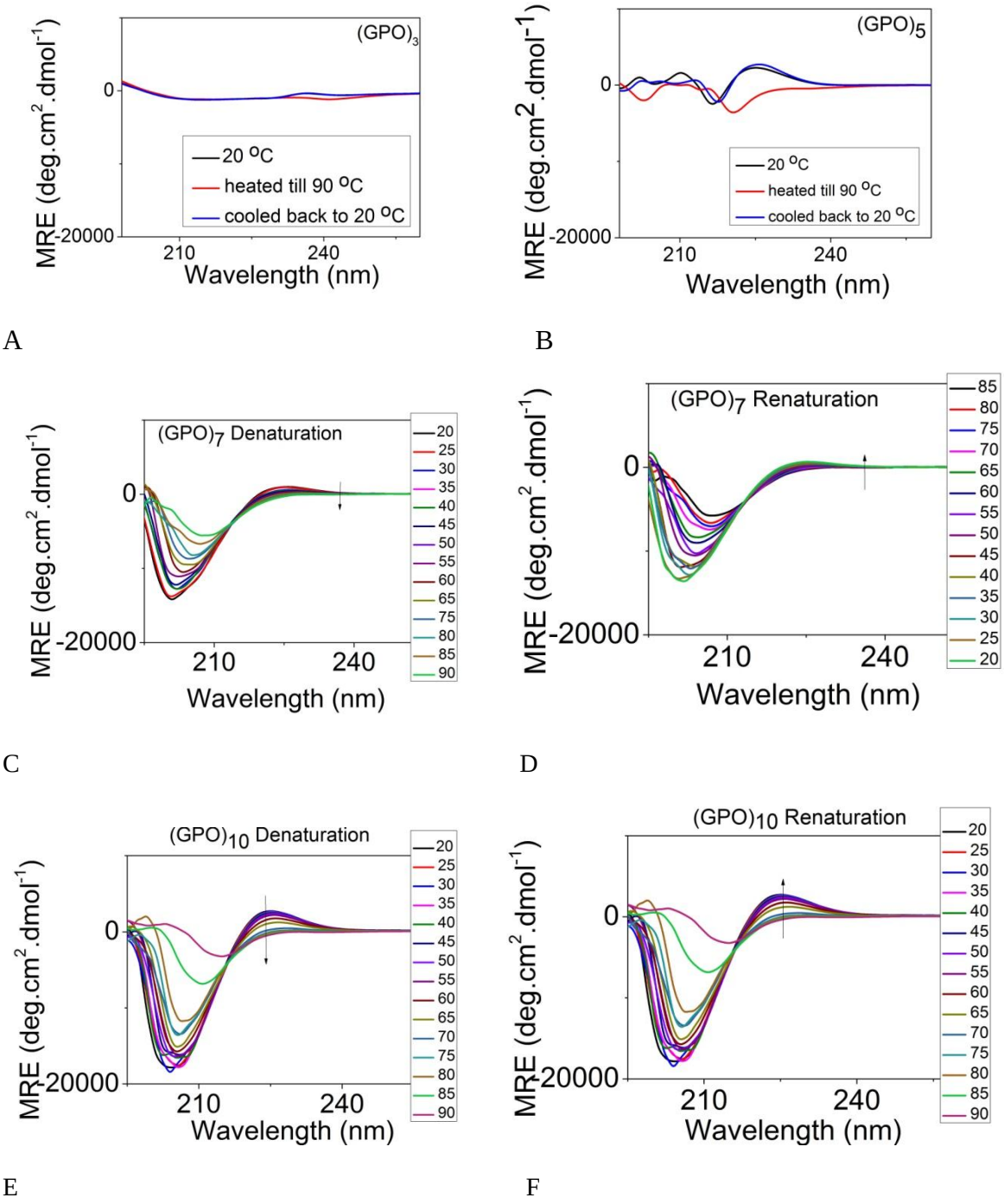


Figure 3.5 CD analysis of melting for $(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ peptides. Samples were at 1 mg/ml concentration, pH-7 at 4 °C and 2 days incubation (A) $(GPO)_3$ automated melting analysis. (B) $(GPO)_5$ automated melting analysis (C) $(GPO)_7$ manual denaturation (D) $(GPO)_7$ manual re-naturation (E) $(GPO)_{10}$ manual denaturation (F) $(GPO)_{10}$ manual re-naturation.

In the denaturation process of the $(GPO)_7$ peptides (Figure 3.5C), the isodichroic point was observed at ~ 205 nm and from the manual melting scans the melting point was determined to be in between 35 °C and 40 °C. The re-naturation point was found to be in-between 40 and 45 °C (Figure 3.5D).

The denaturation of $(GPO)_{10}$ peptides, (Figure 3.5E) exhibits the melting point in-between 65 °C and 70 °C and the isodichroic point at 207 nm. The renaturation point was found to be in-between 65 °C and 70 °C. With temperature decrements, the minimum was again seen to shift to higher wavelengths for both $(GPO)_7$ and $(GPO)_{10}$ peptides (Figure 3.5F).

Corresponding melting curves for the $(GPO)_3$ peptide of MREs at 225 nm are shown in Figure 3.8A (in black). Melting has been clearly established with a gradually decreasing MRE curve. In the same figure, the melting curve of $(GPO)_5$ peptides is shown, with MREs at 225 nm in red. However, the clear melting points could not be specified for either $(GPO)_3$ or $(GPO)_5$ peptides, because their transitions were not cooperative in nature.

Melting curves, MRE at 225 nm for $(GPO)_7$ and $(GPO)_{10}$ peptides are shown in blue and pink respectively in Figure 3.8A. $(GPO)_{10}$ peptides have shown a clear cooperative transition.

3.3.5 Triple helical stabilities and enthalpies- Analysis by DSC

DSC was used to study the effect of varying chain lengths of GPO peptides on melting, enthalpy and heat capacity change.

Differential scanning calorimetry monitors heat flow effects associated with phase transitions and chemical reactions as a function of temperature. It is a thermo-analytical technique in which the difference in amount of heat required to increase the temperature of reference and sample is measured as a function of temperature. Reference and sample are maintained at same temperatures (O'Neill 1964).

For a two state unfolding process, the enthalpy of unfolding (ΔH), the transition temperature (T_m) and the heat capacity change (ΔC_p) upon heating are sufficient thermodynamic parameters to characterize the unfolding process [124].

The relationship between Gibbs free energy (ΔG), ΔH and ΔS can be described as follows,

$$\Delta G(T) = \Delta H(T) - T\Delta S(T)$$

[125]

Equation 3.6

where ΔH and ΔS are the enthalpy and entropy changes of peptide unfolding respectively. The temperature dependence of ΔH and ΔS is defined by the heat capacity change upon unfolding, ΔC_p in Equation 3.7 as follows:

$$\Delta C_p \frac{d\Delta H(T)}{(dT)} = Td\Delta S(T)/(dT)$$

Equation 3.7

If ΔCp is independent of temperature, the temperature dependence of Gibbs free energy gets defined in Equation 3.8

$$\Delta G(T) = \left[T_m - \frac{T}{(T_m)}\right] \times \Delta H(T_m) + \Delta C_p(T - T_m) + T\Delta C_p \ln\left(\frac{T_m}{T}\right)$$

[126]

Equation 3.8

Usually two independent enthalpies can be determined directly from a single DSC experiment i.e. calorimetric enthalpy ΔH_{cal} (T_m) and fitted enthalpy ΔH (T_m) [124, 127, 128].

Melting point/transition temperature analysis was performed. The principal behind DSC analysis is Van't Hoff equation.

$$\frac{d}{dt}(\ln K) = \frac{\Delta H}{RT^2}$$

[129]

Equation 3.9

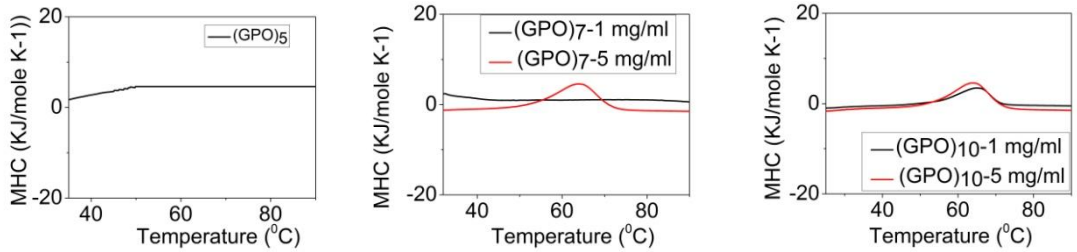
Where K= Equilibrium constant of the process

T= Absolute temperature

ΔH= van't Hoff Enthalpy

In this section, minimum concentration required for triple helix formations and supramolecular assembly formations for GPO peptides have been examined.

Peptides solutions were prepared at 1 mg/ml and 5 mg/ml concentrations in 20 mM PBS buffer (150 mM NaCl) and incubated at 4 °C for 2 days. These concentrations are effective only for the longer chain lengths of peptides like $(GPO)_{7,}$ (Figure 3.6B) and $(GPO)_{10}$ (Figure 3.6C). No unwinding was seen for the shorter length peptides like $(GPO)_3$ and $(GPO)_5$ even by using concentrations of 1 mg/ml or 5 mg/ml. The phenomenon is shown in Figure 3.6A where no triple helical peak for $(GPO)_5$ peptides was observed at 5 mg/ml. Therefore, study on shorter peptides e.g. $(GPO)_3$ had to be abandoned.



A B C

Figure 3.6 DSC thermograms for (GPO) peptide samples, incubated for 2 days, at 4 °C, pH 7 in 20 mM PBS buffer at a scan rate of 1 °C/minute (A) $(GPO)_5$ peptides at concentrations of 5 mg/ml (B) (GPO)7 peptides at 1 and 5 mg/ml concentrations in 20 mM PBS (C) (GPO)10 peptides at 1 and 5 mg/ml concentrations in 20 mM PBS

DSC analysis was performed again with all the peptides in water, at 1 mg/ml concentration, incubated at 4 °C for 2 days. No peaks were observed for the $(GPO)_3$, $(GPO)_5$ and $(GPO)_7$ peptides and a small peak was visible for the $(GPO)_{10}$ peptides. (Figure 3.7A and Figure 3.7B).

The data from the analysis of peptides in 20 mM PBS buffer, in 1 and 5 mg/ml concentrations has been used to determine the melting point and enthalpies.

For $(GPO)_{10}$ peptides, peaks were observed for triple helices in 20 mM PBS buffer pH 7 (see Figure 3.6C) at both concentrations of 1 mg/ml and 5 mg/ml. For 1 mg/ml concentration, the MP was determined to be 64.23 °C and enthalpy was calculated to be 362.16.84 KJ/mol. For 5 mg/ml, the MP was found to be 63.06 °C and enthalpy calculated to be 340.66 KJ/mol (please refer Table 3.3).

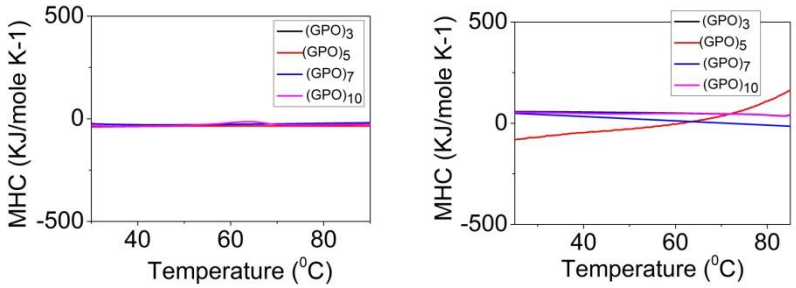


Figure 3.7(A) Heating thermograms in water, $(GPO)_3$, $(GPO)_5$, $(GPO)_7$, $(GPO)_{10}$ peptide at 1 mg/ml (B) Cooling thermograms in water, $(GPO)_3$, $(GPO)_5$, $(GPO)_7$, $(GPO)_{10}$ peptides at 1 mg/ml.

No peaks were observed for the $(GPO)_7$ peptides at a concentration of 1 mg/ml. For the 5 mg/ml sample, a single peak was observed (Figure 3.6B). MP of $(GPO)_7$ triple helices was 33.8 °C and the calculated enthalpy was 187.22 KJ/mol. These were determined from the Van't Hoff equation, through the software, NanoAnalyse (refer Table 3.3).

3.3.6 Effect of salt on the stability of triple helices

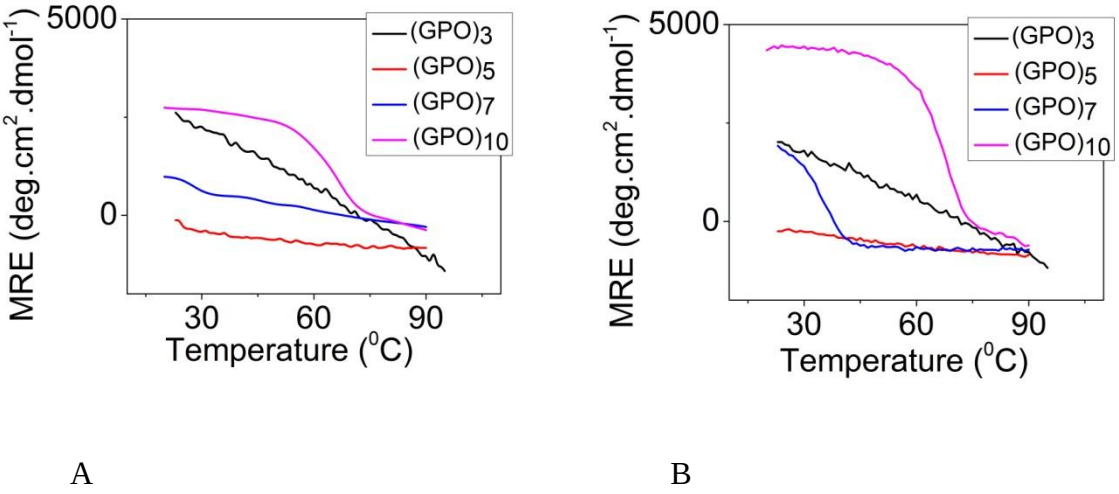


Figure 3.8 Analysis by CD; Impact of salt on stability of GPO peptides of varying chain lengths, depicted by curves for MREs at 225 nm (A) Melting curves in 20 mM PBS buffer (150 mM NaCl salt), 20 to 90 °C (B) Melting curves in 100 mM PBS buffer (213 mM NaCl salt) 20 to 90 °C.

Table 3.3 DSC analysis; Table of enthalpies and transition temperatures for (GPO) peptides with details on error factor (should be ideally less < 1), T= Transition temperature, Enthalpy, ΔS= change in heat capacity, ΔH= change in enthalpy.

Peptides	Buffer	Concentration	Error factor	T (°C)	Enthalpy (KJ/mol)	ΔS KJ/molK^{-1}	ΔH KJ/mol
$(GPO)_{10}$	water	5 mg/ml	0.09	63.52	328.36	0.088	29.795
		1 mg/ml	0.08	63.8	400	0.097	32.762
	20mM PBS pH-7	5 mg/ml	0.09	63.06	340.663	0.100	33.756
		1 mg/ml	0.09	64.23	362.16	0.105	35.67
	Citrate buffer pH 3.5	1 mg/ml	0.03	65.18	434.18	0.031	10.503
	Tris buffer pH 8.0	1 mg/ml	0.03	64.5	452.23	0.043	14.709
$(GPO)_7$	20mM PBS pH-7	5 mg/ml	0.2	33.82	187.22	0.125	38.43
		1 mg/ml	-	No peak	-	-	-
	Citrate buffer pH 3.5	1 mg/ml	-	No peak	-	-	-
	Tris buffer pH 8.0	1 mg/ml	-	No peak	-	-	-

GPO peptide solutions of varying chain lengths, $(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ were prepared in 100 mM PBS buffer (213 mM NaCl), pH 7 and incubation time of 2 days at 4 °C (B). The usual buffers i.e. 20 mM PBS with salt contents of 150 mM NaCl are also commonly used for the stability studies (Figure 3.8A).

Table 3.4 Analysis by CD; Impact of salt on melting of GPO peptides of varying chain lengths; MREs at 225 nm from Figure 3.8.

Peptides	Ellipticity (MRE) at 20 °C in 20 mM PBS buffer (150 mM NaCl)	Ellipticty (MRE) at 20 °C in 100 mM PBS buffer (213 mM NaCl)
$(GPO)_3$	2100	2100
$(GPO)_5$	-	-
$(GPO)_7$	1000	2000
$(GPO)_{10}$	2800	4100

From the melting analysis curves (MRE at 225 nm) of all peptides, it was observed that the high salt environment provides more support for the triple helices and for the helical meltings as compared to a low salt environment (see Table 3.4).

Cooperative transitions were observed only on $(GPO)_{10}$ peptides in 20 mM PBS buffer (Figure 3.8A) in a low salt environment. In a high salt environment (Figure 3.8B) cooperative transitions were observed on both $(GPO)_7$ and $(GPO)_{10}$ peptides; $(GPO)_5$ and $(GPO)_3$ peptides were too short to undergo any cooperative transitions, even in a high salt environments.

3.3.7 Effect of pH on the stability of triple helices;

After performing the aforementioned analyses, $(GPO)_{10}$ and $(GPO)_7$ peptides were picked to study the secondary structural changes in triple helices with gradual temperature change (increments and decrements) at various pHs. Melting analysis was also performed. Detailed melting experiments for all peptides have already been performed at pH 7, in the above sections (with 20 mM PBS buffer, pH7). Stabilities of triple helices have been determined here, at the two different pHs of opposite ends, pH 3.5 and pH 8 (Table 3.5).

Table 3.5 CD analysis; Effect of pH on triple helical stabilities; $(GPO)_7$ and $(GPO)_{10}$ peptides

pH	Peptides	Melting point	Re-naturation point
pH 3.5	$(GPO)_7$	40 °C	35 °C
	$(GPO)_{10}$	65- 70 °C	50 °C
PH 8	$(GPO)_7$	- (noisy)	- (noisy)
	$(GPO)_{10}$	65 °C	45 °C

3.3.7.1 $(GPO)_{10}$ peptides

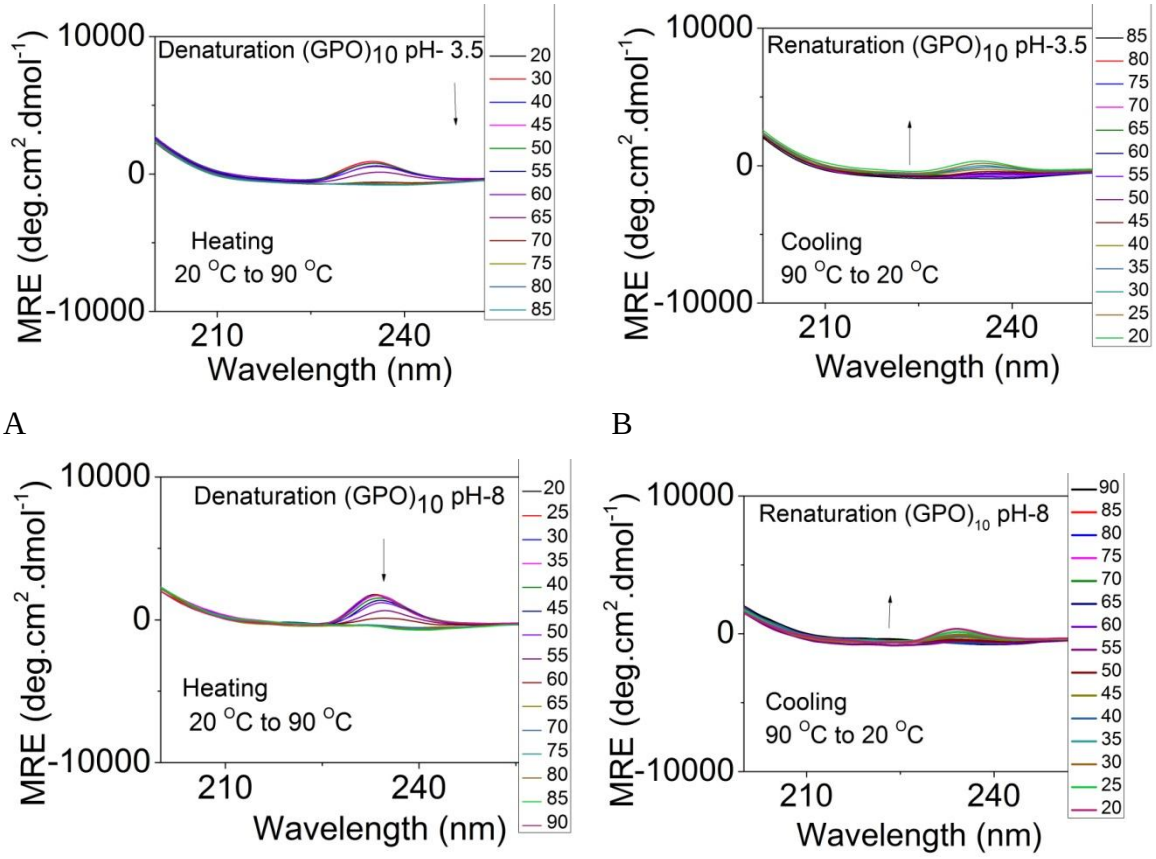
In Figure 3.9, at pH 3.5, positive CD peaks (maxima) at 235 nm were seen. These were the triple helices. With the application of temperature increments (Figure 3.9A), the MP was determined to occur in-between 65 °C and 70 °C. The re-naturation point was ~ 50 °C (Figure 3.9B).

At pH 8, the $(GPO)_{10}$ peptide exhibits triple helices formation. On heating (Figure 3.9C), the structure melts at ~ 65 °C and again re-natures at around 45 °C (Figure 3.9D).

3.3.7.2 $(GPO)_7$ peptides

For $(GPO)_7$ peptide at pH 3.5, the triple helices unfold and re-fold with cycles of heating and cooling as shown in (Figure 3.10A) and (Figure 3.10B) respectively. Random coils are shown as the minima (negative dip) at around 200 nm. These shift to higher and lower wavelengths respectively, with gradual temperature increments and decrements. A melting point was observed at ~ 40 °C and the re-naturation point at 35 °C.

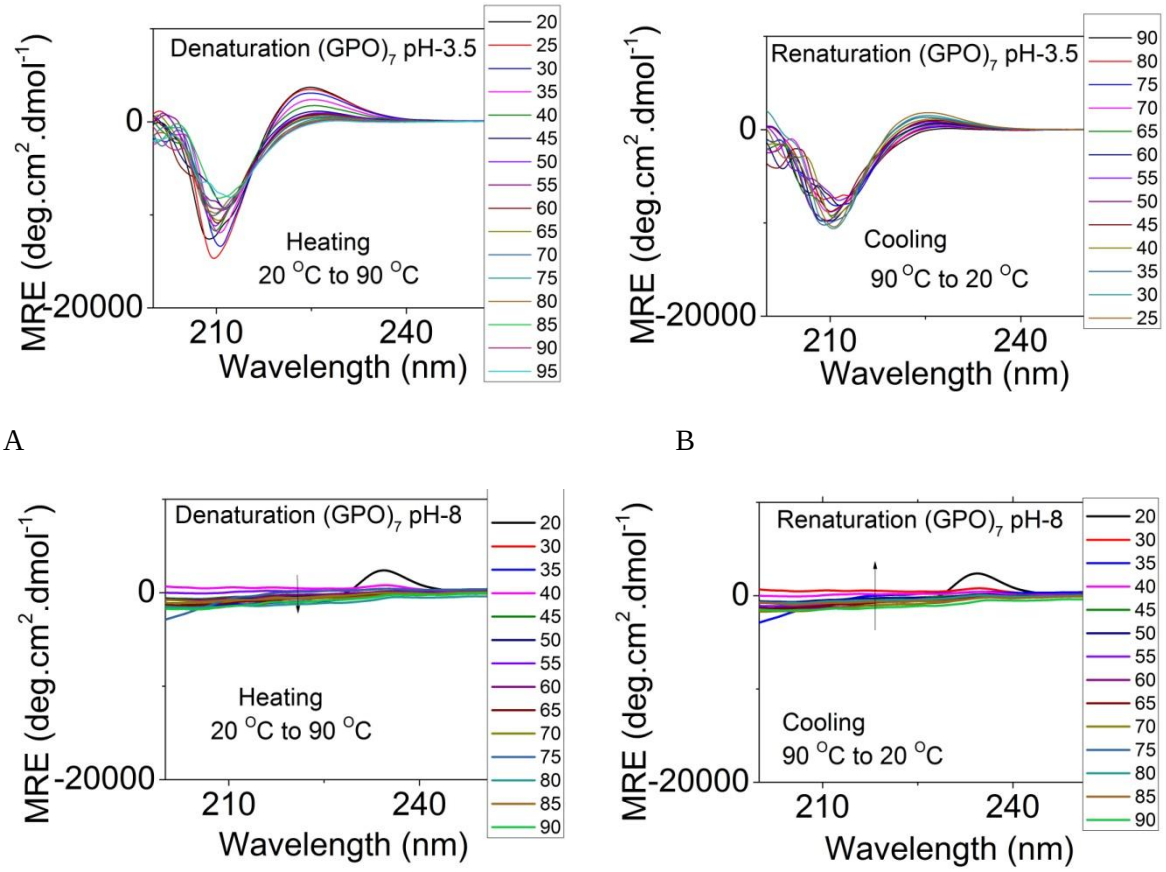
At pH 8, the melting point and renaturation point were difficult to determine because the spectra were noisy, (see Figure 3.10C and Figure 3.10D). Based on the activity of $(GPO)_{10}$ triple helices, where re-naturation point of the helices was achieved at ~ 40 °C at pH 8 and it was achieved at ~ 50 °C with pH 3.5, the pH which favored triple helical stabilities the most after the neutral pH, was pH 3.5.



A B

C D

Figure 3.9 CD stability studies at various pH; $(GPO)_{10}$ peptides (A) $(GPO)_{10}$ at pH-3.5, temperature scan of 20 °C to 90 °C (B) $(GPO)_{10}$ at pH-3.5, temperature scan of 90-20 °C; Species formed with $(GP0)_{10}$ with denaturation and re-naturation seemed unclear (C) $(GPO)_{10}$ at pH-8, temperature scan of 20 °C to 90 °C (D) $(GPO)_{10}$ at pH-8, temperature scan of 90-20 °C.



A B

C D

Figure 3.10 CD stability studies at various pH; $(GPO)_7$ peptides (A) $(GPO)_7$ at pH- 3.5, temperature scan of 20 °C - 90 °C (B) $(GPO)_7$ at pH- 3.5, temperature scan of 90 °C - 20 °C (C) $(GPO)_7$ at pH- 8, temperature scan of 20 °C – 90 °C (D) $(GPO)_7$ at pH- 8, temperature scan of 90 °C - 20 °C.

3.3.8 Differences in stabilities for $(GPO)_7$ peptides and amidated $(GPO)_7$ peptides

Manual melting analysis was performed to study the differences in the stabilities of the $(GPO)_7$ peptide and an amidated $(GPO)_7$ peptide. From Figure 3.11, it was seen that the amidated peptide, NH_2- $(GPO)_7$-NH_2 has an MP between 60 °C and 65 °C (Figure 3.11C); The MP of $(GPO)_7$ triple helices was seen between 35 and 40 °C (Figure 3.11A), see Table 3.6. It was believed that the amidation of the end terminals add a distinctive stability to the $(GPO)_7$ helices. It was observed that the isodichroic point was ~205 nm for the amidated peptides. This was same for the $(GPO)_7$ peptides.

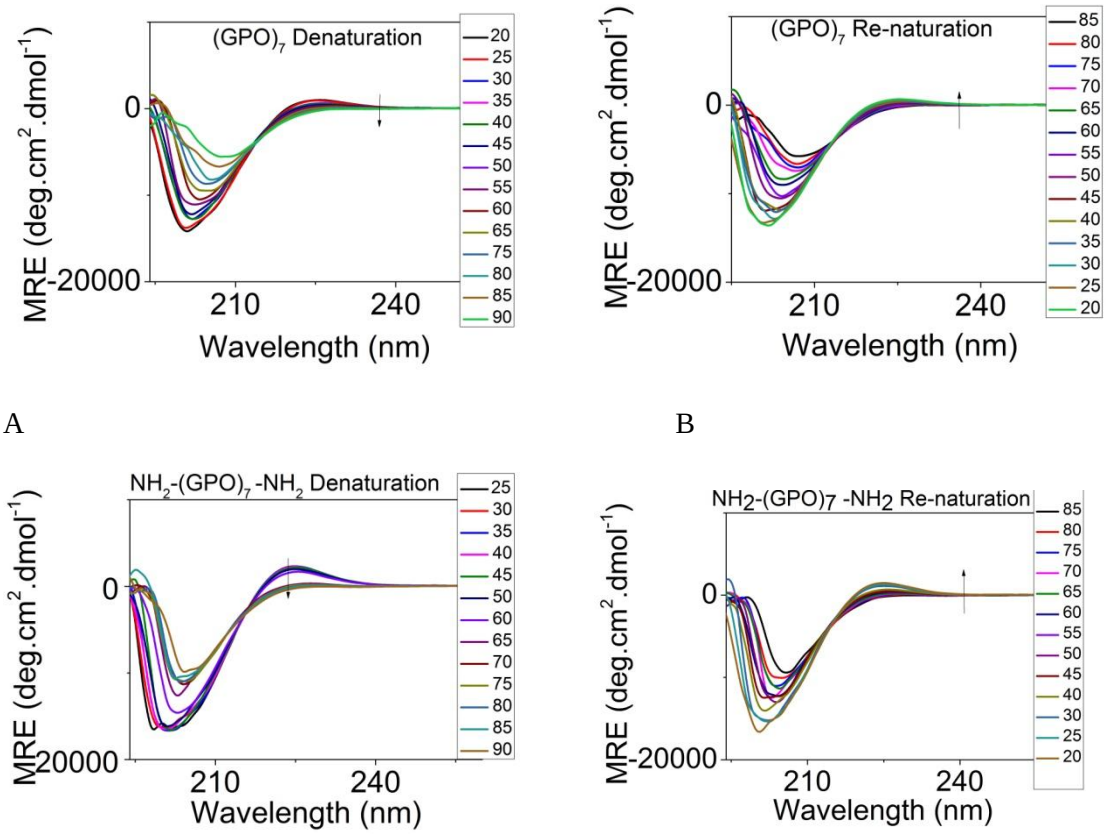


Figure 3.11 CD analysis; Stability differences between $(GPO)_7$ and amidated $(GPO)_7$ peptides (A) $(GPO)_7$ manual melting, denaturation (B) $(GPO)_7$ renaturation (C) NH_2-$(GPO)_7$-NH_2 manual melting, denaturation (D) NH_2-$(GPO)_7$-NH_2 renaturation

Table 3.6 Analysis by CD; Melting point and isodichroic point analysis for $(GPO)_7$ and amidated $(GPO)_7$ peptides

Peptides	Melting temperature	Isodichroic point
NH_2- $(GPO)_7$-NH_2	60 to 65 °C	~205 nm
NH_2-$(GPO)_7$-COOH	35 to 40 °C	~205 nm

For amidated $(GPO)_7$ peptides, maxima for helices remained intact at 225 nm. Random coil formation was seen with the shifting of the minima at ~ 205-210 nm to higher and lower wavelengths, throughout the denaturation (Figure 3.11C) and the re-naturation process (Figure 3.11D). For amidated peptides, the minima shifts at ~ 205 nm was small as compared to $(GPO)_7$ peptides. Therefore, higher stability of the amidated $(GPO)_7$ triple helices was recorded and confirmed.

MRE curves have been shown at 225 nm for the melting process, with subsequent temperature increment scans, for both the peptides in Figure 3.12.

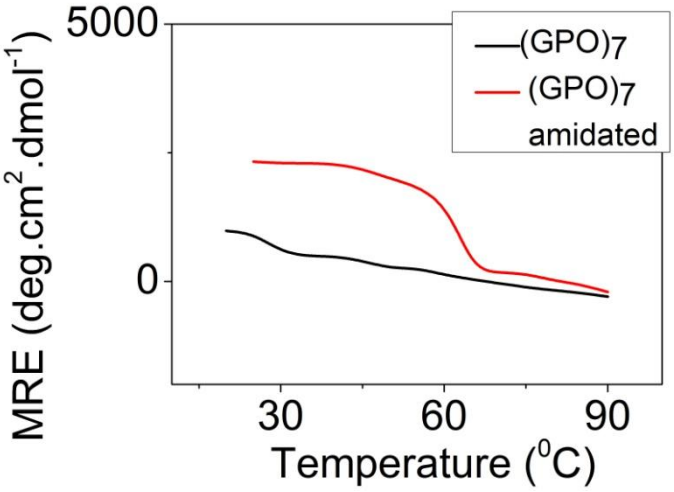


Figure 3.12 Analysis by CD; $(GPO)_7$ and amidated $(GPO)_7$ peptides. MRE melting plots at 225 nm, temperature scan 20 °C to 90 °C in 20 mM PBS high salt buffer (150 mM NaCl salt)

3.4 Supramolecular assemblies

Transmission electron microscopy, dynamic light scattering and atomic force microscopy have been used to characterize the (GPO) supramolecular assembly formations.

3.4.1 Dynamic light scattering analysis of the supramolecular assemblies

3.4.1.1 Critical concentration determination for supra assembly formations

The structures formed with the varying chain lengths of (GPO) peptides were successfully differentiated via DLS. The technique was used to derive translational diffusion coefficient of the samples, which were processed to determine the size distribution of the assemblies, in terms of hydrodynamic diameter (nm).

Peptide samples of $(GPO)_3$, $(GPO)_5$, $(GPO)_7$, $(GPO)_{10}$ were prepared at 0.2, 0.5 and 1 mg/ml concentrations. These were prepared in 20 mM PBS buffer at pH 7 and incubated at 4 °C for 3 days.

For $(GPO)_3$ peptides prepared at 0.2 mg/ml concentrations (Figure 3.13A), assemblies of hydrodynamic diameter 141.77 nm were observed, which were possible random aggregations. At a concentration of 0.5 mg/ml (Figure 3.14A) two kinds of assembly species - the 1.7 nm species (the triple helices) and the 615 nm (aggregates) were seen.

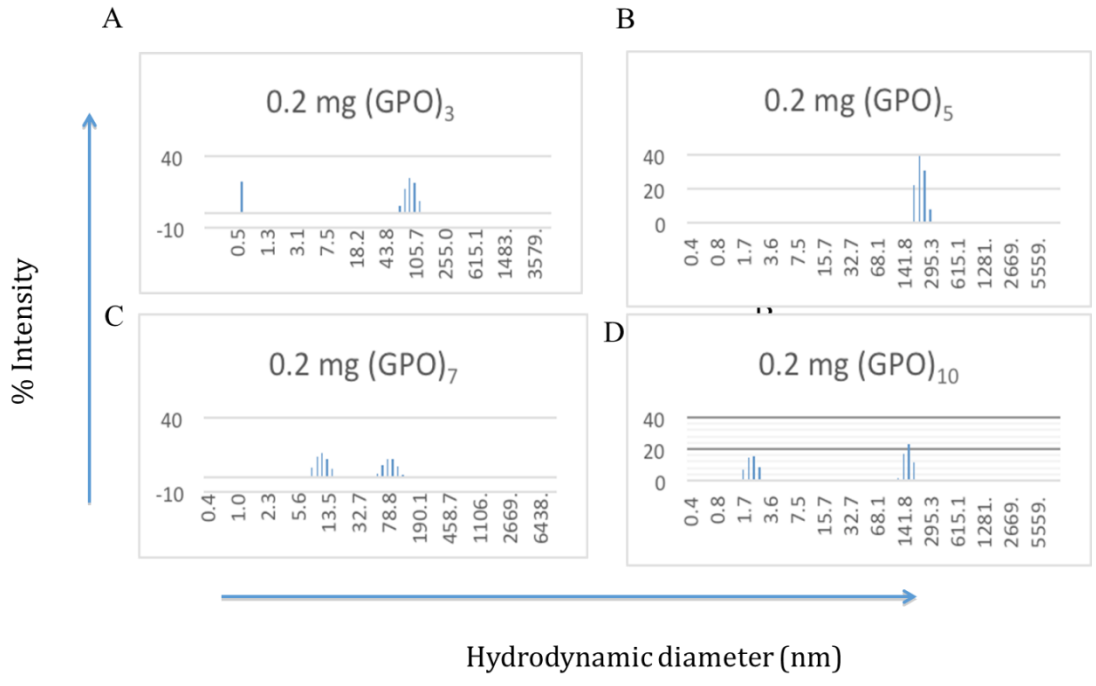


Figure 3.13 DLS analysis (A) $(GPO)_3$ at 0.2 mg/ml concentration (B) $(GPO)_5$ at 0.2 mg/ml concentration (C) $(GPO)_7$ at 0.2 mg/ml concentration and (D) $(GPO)_{10}$ at 0.2 mg/ml concentration

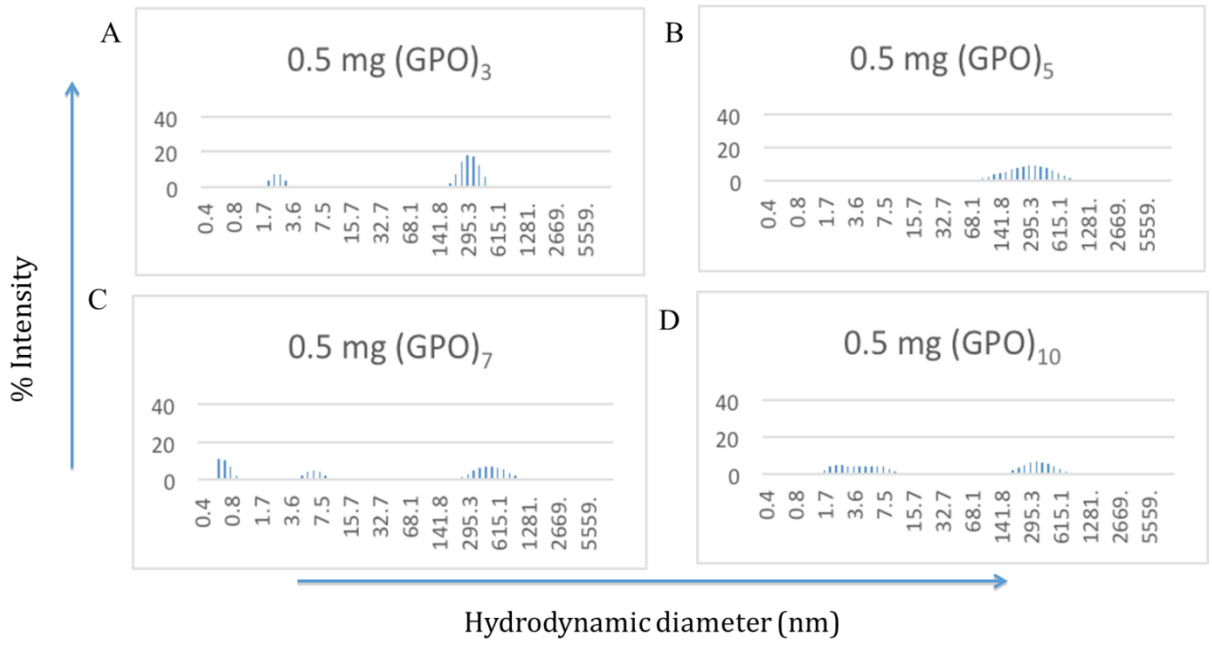


Figure 3.14 DLS analysis $(GPO)_3$ at 0.5 mg/ml concentration (B) $(GPO)_5$ at 0.5 mg/ml concentration (C) $(GPO)_7$ at 0.5 mg/ml concentration and (D) $(GPO)_{10}$ at 0.5 mg/ml concentration

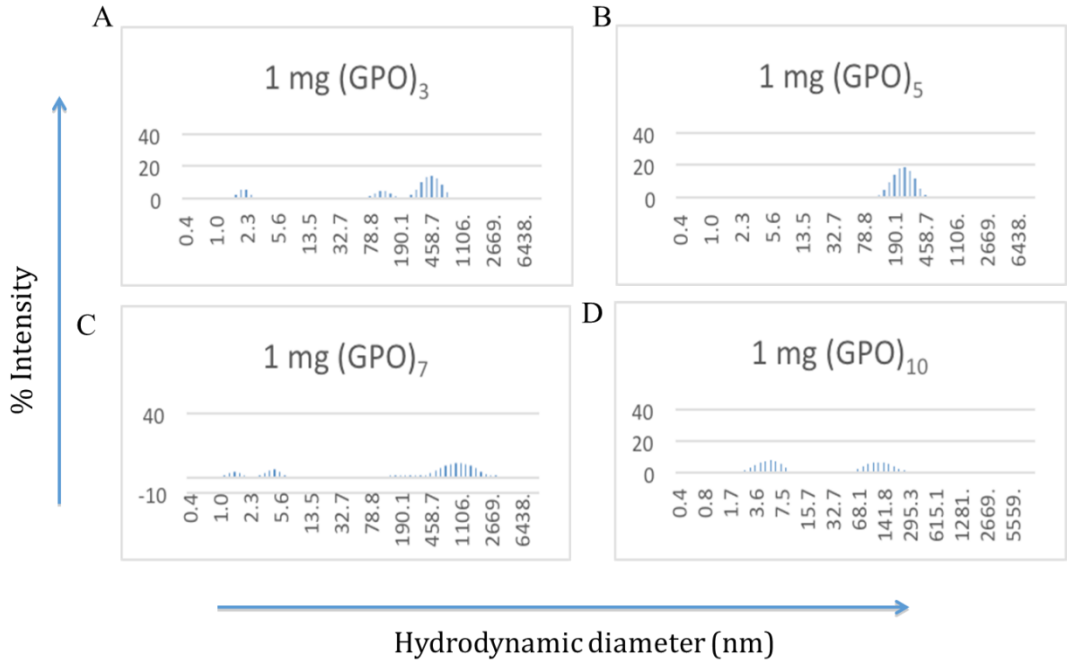


Figure 3.15 DLS analysis (A) $(GPO)_3$ at 1 mg/ml concentration (B) $(GPO)_5$ at 1 mg/ml concentration (C) $(GPO)_7$ at 1 mg/ml concentration and (D) $(GPO)_{10}$ at 1 mg/ml concentration

At 1 mg/ml (Figure 3.15A), three assemblies of hydrodynamic diameter, 1.7 nm (triple helices), 141.77 nm (aggregates) and 615 nm (aggregates) were observed. A minimum conc. of 0.5 mg/ml and above seemed essential to visualize $(GPO)_3$ helices.

For $(GPO)_5$ peptide, at all three concentrations of 0.2 mg/ml (Figure 3.13B), 0.5 (Figure 3.14B) and 1 mg/ml (Figure 3.15B), only aggregates of ~ 615 nm diameter were seen. It was speculated that the helices could have fused together to form aggregates, at all concentrations.

For longer peptides like $(GPO)_7$ peptides, right from 0.2 mg/ml (Figure 3.13C) both species, the triple helices (1.7 nm) and higher order assemblies (~141.7 nm) were observed. At concentrations of 0.5 mg/ml (Figure 3.14C) and 1 mg/ml (Figure 3.15C), species bigger than 615 nm and 1000 nm diameter were also seen.

For the longest $(GPO)_{10}$ peptides, both aggregates and helices formed at all three concentrations (Figure 3.13D, Figure 3.14D, Figure 3.15D). At all three concentration, the 615 nm diameter species were seen, along with the 1.7 nm species (triple helices).

DLS study indicated that peptides of longer chain lengths like $(GPO)_7$ and $(GPO)_{10}$ peptides gave rise to well developed unique triple helical assemblies, along with associated supramolecular assemblies right from low concentrations of 0.2 mg/ml until 1 mg/ml. The shorter peptides like $(GPO)_5$ and $(GPO)_3$ showed only associated species at all concentrations.

3.4.2 TEM analysis of the supramolecular assemblies

3.4.2.1 Critical concentration determination for supramolecular assemblies

$(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ peptide solutions were prepared at 1 % (1 mg in 100 µl), 5 %, 10 % and 20 % concentrations (weight by volume). All peptide samples were prepared in both 100 mM PBS buffer (213 mM NaCl) and in water.

Pasini et al., [130] have heavily used weight % units for their experimental work involving higher order assemblies, like gels. In triple helical stability studies by Brodsky et al., concentrations in terms of milligrams per milliliters (weight % units) were also used, for

comparing the stabilities of (POG)n and (PPG)n collagen mimetic peptides [120]. Weight % units have also been used in this thesis.

Previous studies of Kar et al. have been referred, to understand and establish the criteria of the supramolecular assembly formation [48]. To observe peptide supra-molecular assemblies, buffers as controls have been used throughout this research.

Experiments have been performed with shortest and the longest GPO peptides. $(GPO)_3$ peptides were prepared in 1 % (1 mg in 100 µl), 5% and 10% concentrations in both 100 mM PBS buffer and water. Both were incubated at 4 °C and 37 °C. The incubations were checked at 1 day, 1 week, 2 weeks and 1 month of incubation time periods, to examine the supramolecular assembly formations.

Supramolecular self-assemblies were observed at 20 % concentrations, when incubated at temperatures of both 4 °C and 37 °C for $(GPO)_3$ peptides at incubation time of 2 weeks (see Figure 3.16A and Figure 3.16B). No assemblies were observed for 1 %, 5 % and 10 % concentrations.

$(GPO)_{10}$ peptides self-associated to form supramolecular assemblies in the100 mM PBS buffer at 1% concentration, at 1 week incubation time and at 4 °C (see Figure 3.16C). The $(GPO)_{10}$ peptide also formed supramolecular assemblies at 37 °C at 1% concentration when incubation time was 1 week (see Figure 3.16D). These peptides have been further analysed with atomic force microscopy for structural details.

The $(GPO)_5$ peptides were seen to self associate to form supramolecular assemblies at concentrations of 5 %, at incubation temperatures of 4 °C, in water (Figure 3.17A). Fibrillar structures for $(GPO)_5$ peptides were seen at 5 % concentration, at 37 °C (Figure 3.17B). Globular assemblies were also noticed at concentrations of 20 % for $(GPO)_5$ peptides, at 4 °C, in water (see Figure 3.17C). For $(GPO)_7$ peptides, at a concentration of 5 % at 4 °C, supramolecular assemblies were seen in water (Figure 3.17D).

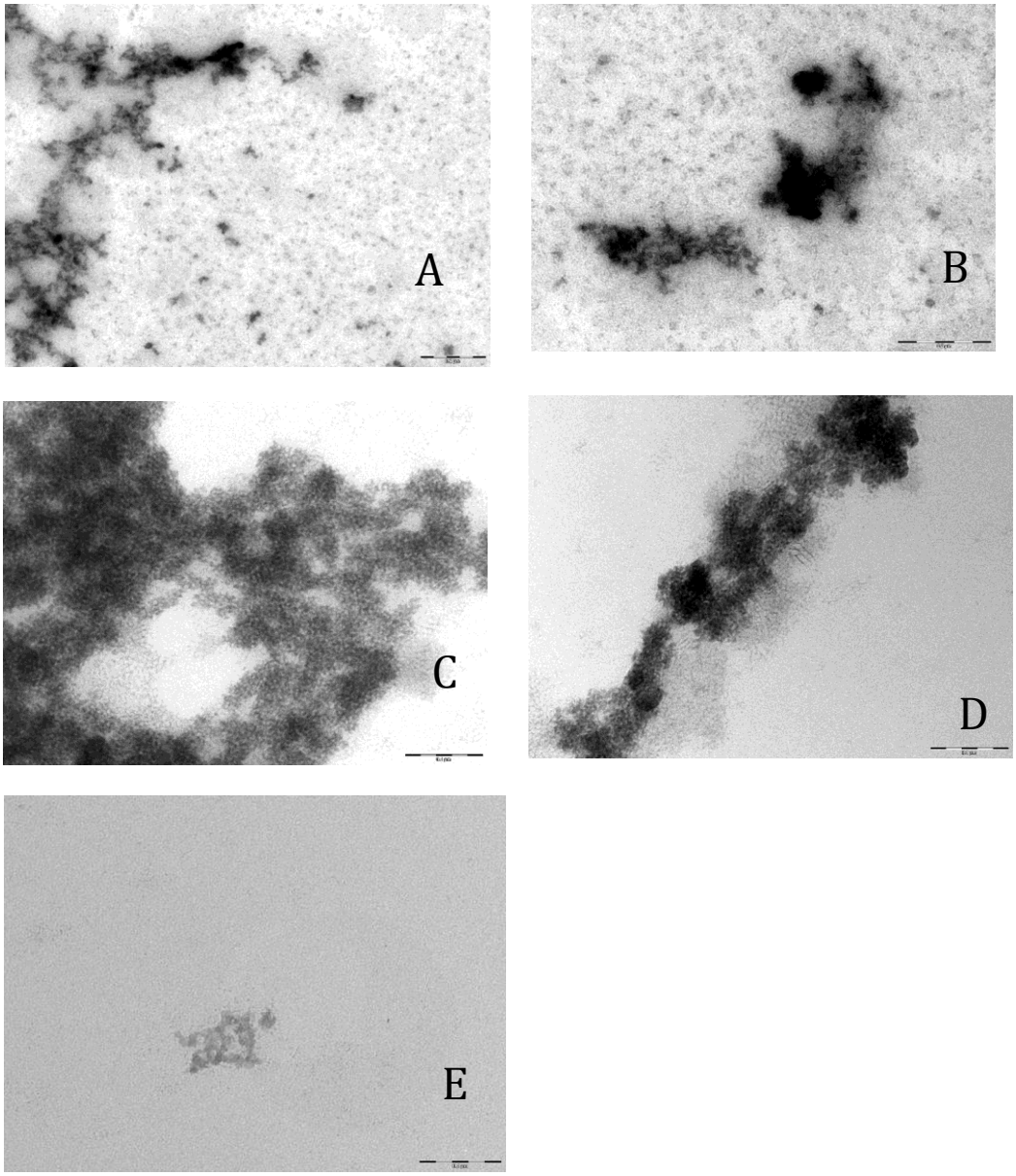


Figure 3.16 TEM analysis; $(GPO)_3$ and $(GPO)_{10}$ peptides in 100 mM PBS buffer (A) $(GPO)_3$ formed supramolecular assemblies at 20 % concentration and 2 weeks incubation at 4 °C; scale was 200 nm (B) $(GPO)_3$ showed supramolecular assemblies at 20 % concentration, 2 weeks incubation, at 37 °C; scale was 200 nm (C) $(GPO)_{10}$ showed supramolecular assemblies at concentration of 1 %, 1 week incubation at 4 °C; scale was 100 nm. (D) $(GPO)_{10}$ showed supramolecular assemblies at a concentration of 1%, 1 week incubation, at 37 °C; scale was 200 nm (E) control, 100 mM PBS buffer. TEM artefacts were differentiated from supramolecular assemblies based on following references [131]**,** [132]**.**

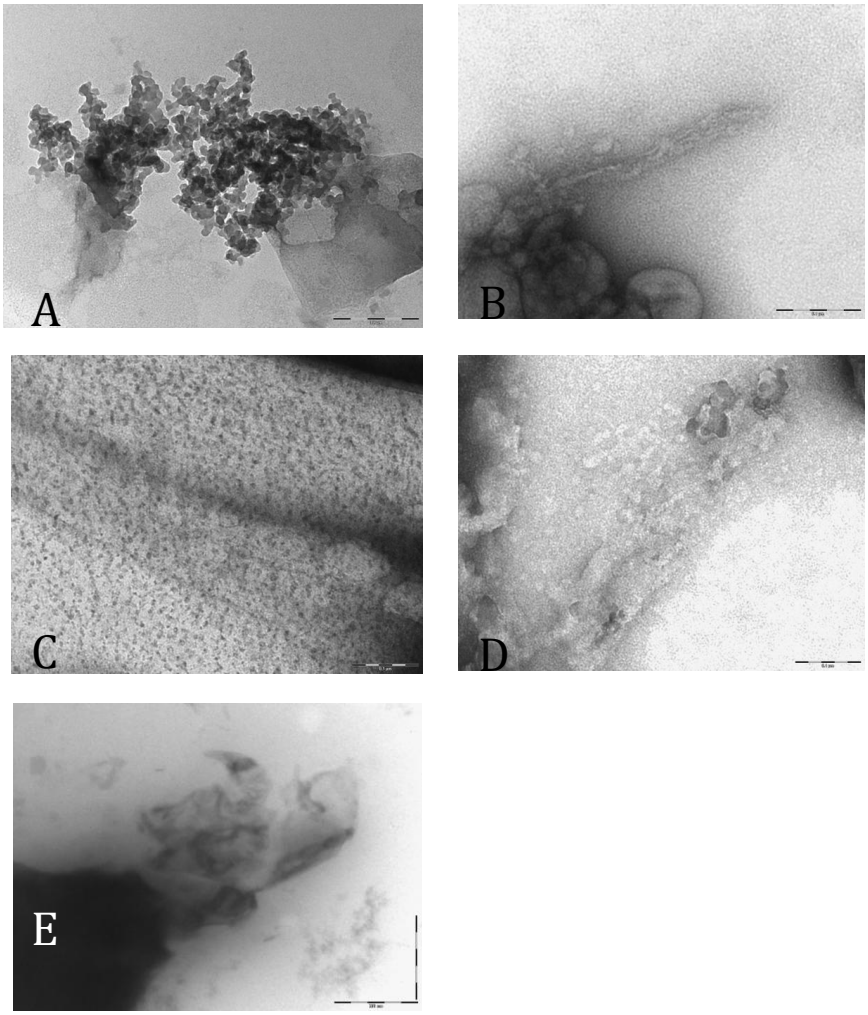


Figure 3.17 $(GPO)_5$ and $(GPO)_7$ peptides in water; TEM analysis (A) $(GPO)_5$ at 5 % concentration, at 4 °C (B) $(GPO)_5$ at 5 % concentration, at 37 °C (C) $(GPO)_5$ at 20 % concentration, at 4 °C (D) $(GPO)_7$ at 5 % concentration, at 4 °C. (E) control. TEM artefacts were differentiated from supramolecular assemblies based on following references [131], [132].

From literature, turbidity studies on $(POG)_{10}$ have reported the formation of self-associated structures from 1st stage assembly of triple helices [48]. $(POG)_7$ peptides have shown no indication of self association at 7 mg/ml at any temperature studied, while $(POG)_8$ has shown a small rise in turbidity (self-association) at 7 mg/ml at 48 °C. $(POG)_{10}$ peptides have shown rapid formation of aggregates at various temperatures (especially at 7 mg/ml 58 °C, while $(POG)_{12}$ peptides show a high propensity of aggregation even at low concentrations and low temperatures, than seen for $(POG)_{10}$ peptides [48]. Kar et al. reported that for $(POG)_{10}$ peptides, there was no aggregation even after incubation for 1 week at high concentrations of 14 mg/ml at 1 °C - 4 °C, below its melting point [133].

In this chapter, it has been observed that the $(GPO)_{10}$ peptides achieved self-associated supramolecular assemblies at high salt environment, 100 mM PBS buffer (213mM

NaCl) at 1% concentration (10 mg/ml) at 4 °C with 1 week of incubation time. $(GPO)_{10}$ peptides have also formed assemblies at 37 °C, at 1% concentration (10 mg/ml), when incubated for a week.

3.4.3 Effect of pH on supramolecular assembly formations

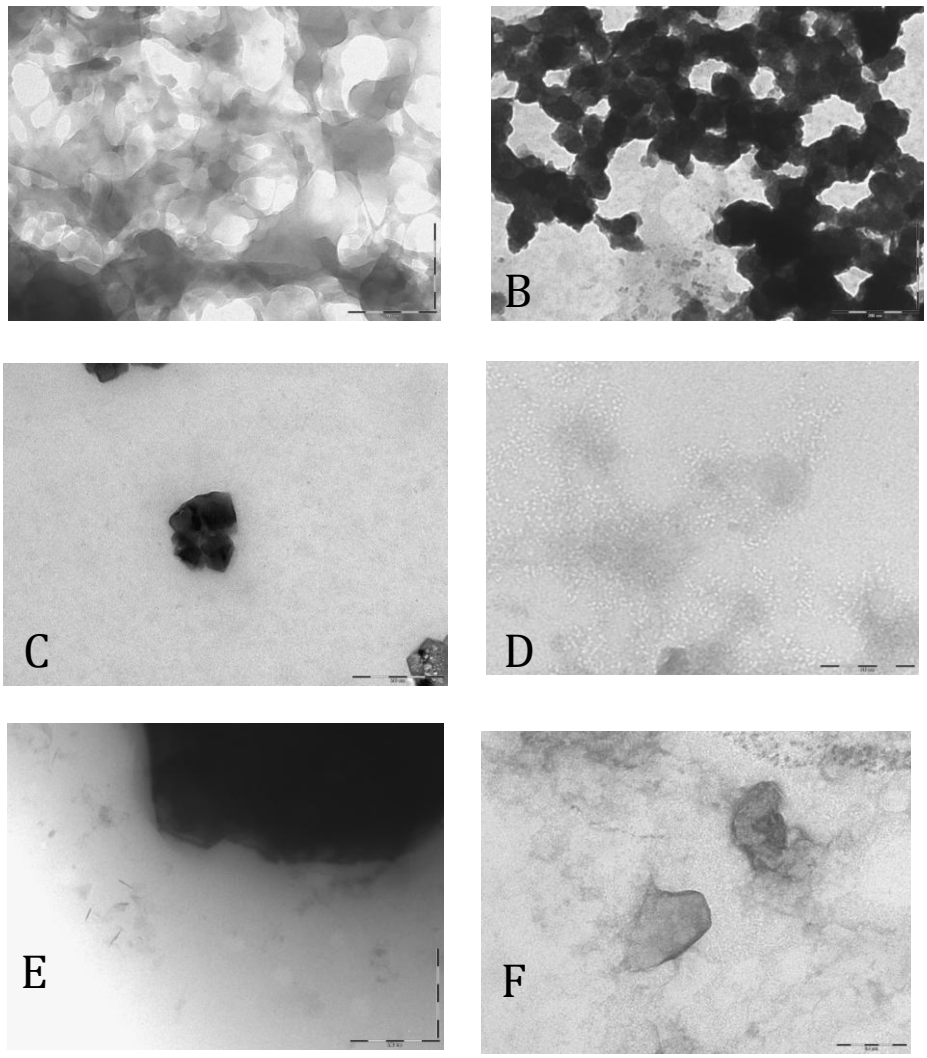


Figure 3.18 TEM analysis (A) $(GPO)_3$ in pH 3.5, scale 100 nm (B) $(GPO)_{10}$ in pH 3.5, scale 200 nm (C) $(GPO)_3$ in pH 8.0, scale 100 nm (D) $(GPO)_{10}$ in pH 8.0, scale pH 8.0. (E) control buffer at pH-3.5 (showed some blobs) (F) control buffer at- pH 8. TEM artefacts were differentiated from supramolecular assemblies based on following references [131], [132].

From the literature, it is known that artificial proteins have been designed which can undergo self-assembly to form hydrogels responsive to pH and other environmental changes [134].

To study the effect of pH on supramolecular assembly formation, once again the shortest and the longest GPO peptides i.e $(GPO)_3$ and $(GPO)_{10}$ peptides were selected. At pH 3.5, in Figure 3.18A, $(GPO)_3$ peptides show unclear assemblies (at scale of 100 nm) and $(GPO)_{10}$ peptides show globular conglomerates in Figure 3.18B at a scale of 200 nm. At pH 8, Figure 3.18C, no noticeable features were seen with $(GPO)_3$ peptides (scaled at 100 nm). With $(GPO)_{10}$ peptides, in Figure 3.18D, non-conclusive structures were recorded (scale was 100 nm). It is speculated that pH 3.5 favors the supramolecular assembly formations more.

3.4.4 $(GPO)_{10}$ supramolecular assemblies, characterisation of morphologies by AFM

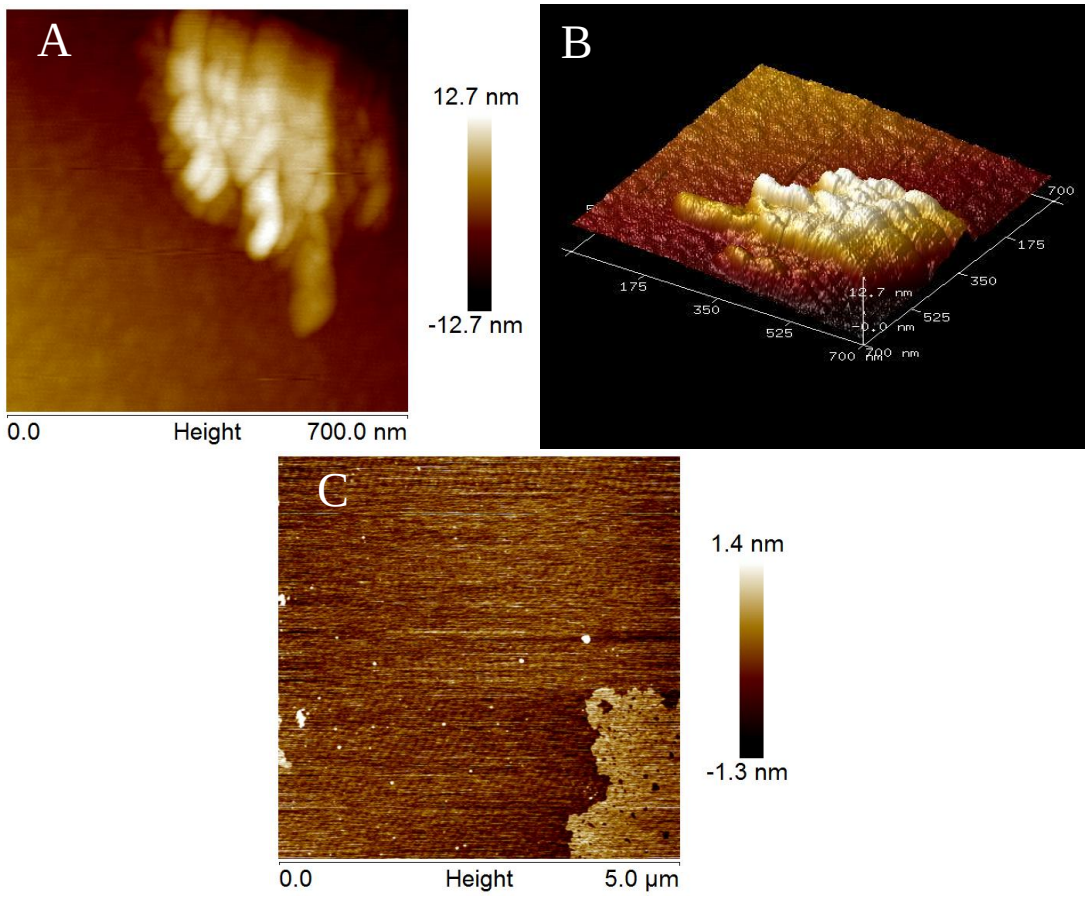


Figure 3.19 AFM imaging; Supramolecular assemblies of $(GPO)_{10}$ peptides; (A) self-associated helical convolutions were visible (B) 3D imaging of supramolecular assembly; all AFM measurements have been performed with NanoScope Analysis. (C) Control as 20 mM PBS buffer. These images have not been reproduced. AFM artefacts have been recognized and avoided by referring [135].

Sample preparations for AFM require low concentrations of peptides ~ 0.01 mg/ml to visualize triple helices in associated forms. Protocols from Tekinay et al. were referred to and modified, for the AFM sample preparations [136]. Samples were prepared on mica and

overnight incubations were carried out at 4 °C. After drying, tapping mode AFM was performed on them within the next day [136]. Scan size was fixed at 700 nm.

Slight broadening for fibrils in the AFM images was observed. This could have occurred due to convolution effects, an effect that would arise due to the finite width of the AFM tip. Effects arising from mechanical and electrostatic contributions of the AFM imaging process tend to add up to small deviations from the original measurements [137].

For $(GPO)_{10}$ peptides, the heights of $(GPO)_{10}$ branched filamentous assemblies were measured to be ~12.7 nm. The widths of the fibrillar convolutions were in-between 10 nm and 40 nm each, approximately. These seemed to assemble together giving rise to larger aggregates (Figure 3.19B). The measurements were performed by NanoScope Analysis.

3.5 Summary of findings

3.5.1 GPO triple helix formations

From CD experiments, it was observed that the triple helical content, indicated by the mean residual ellipticity (MREs) increases with the increasing chain lengths of peptides for all peptides. It is confirmed that the helix formation is directly proportional to chain length, concentration of the mixture and time of incubation.

3.5.2 GPO helical stability studies

Melting point and re-naturation studies indicate that the longer chain length of peptides exhibit higher helical stability i.e. the most helical stability is observed for $(GPO)_{10}$ peptides, followed by $(GPO)_7$ peptides, $(GPO)_5$ and $(GPO)_3$ peptides. It is confirmed that triple helical nature is reversible upon reversal of experimental conditions. Chain length of GPO peptides was found to be inversely proportional to critical concentration for helical unwinding from DSC studies.

It was also observed that amidation of both end terminals of $(GPO)_7$ peptides increases the stability of $(GPO)_7$ helices.

3.5.3 Effect of salt on self-assembly

The triple helical content of the GPO peptides in high salt conditions (213 mM NaCl) was higher as compared to a normal salt environments (150 mM NaCl). With CD melting studies,

cooperative transitions were observed, but only for $(GPO)_{10}$ peptides. In a high salt environment, cooperative transitions were observed in both $(GPO)_7$ and $(GPO)_{10}$ peptides.

3.5.4 Effect of pH on self-assembly

Maximum triple helical content was seen at pH 7 for most of the peptides, followed by pH 3.5 and then pH 8. Following results from DSC and manual CD melting studies, after pH 7, pH 3.5 favored the triple helical stability of helices the most, followed by pH 8. pH 3.5 also supported the formation of supramolecular assembly for (GPO) peptides more than the pH 8.

3.5.5 Self-association kinetics of GPO peptides

From DLS experiments, peptides of longer chain lengths like $(GPO)_7$ and $(GPO)_{10}$ gave rise to well developed unique triple helical assemblies, along with associated structures right from low concentrations (0.2 mg.ml), unlike shorter peptides like $(GPO)_5$ and $(GPO)_3$.

3.5.6 Self-associations of helices into supramolecular assemblies

Self-association was seen to limit by the low concentration of monomers. Manual melting studies with circular dichroism indicate that in the $(GPO)_7$ and $(GPO)_{10}$ peptides, the minima depicting random coils that shift to higher wavelengths (with temperature increments) and later shift to lower wavelengths (with temperature decrements), could result from the breaking of H-bonding and reformation of H-bonding respectively.

4 Effect of temperature cycles on the self assembly of collagen peptides

The effect of temperature cycles on the supramolecular self assemblies of collagen model peptides have been studied in this chapter.

Infrared (IR) spectroscopy has been used to study the changes brought about in the underlying polyproline II chains of the supramolecular assemblies of GPO peptides and the intra and inter hydrogen bonding patterns.

4.1 Background

Previous researchers have investigated factors that might impact on the association of collagen mimetic peptides into disordered aggregates. Among other factors hydration has been proved to be a major driving force for collagen self-assembly [138, 139].

Kar et al., have shown that optimization of peptide length, concentration, temperature etc have an effect on self-assembly. They have also shown that the $(GPO)_{10}$ peptides result in highly branched filamentous assemblies [48, 140]. Fibril formation has been noticed to have a nucleation mechanism, which increases with temperature. The process was noted to be the fastest at a temperature just below melting temperature which indicates that the micro unfolding close to melting drove the supramolecular assembly formations [141]. Kar et al. have shown that supramolecular assemblies are favored at a temperature just below the melting point (MP) of the triple helice [48, 140].

From previous chapters, some information on the MPs of the GPO collagen peptides was collected and based on those, appropriate temperature ranges were picked, to investigate self-assembling properties of the peptides. Small angle X-ray scattering has also been used to investigate the self association properties of the GPO peptides as liquid crystalline arrays.

4.2 Introduction

Infrared spectroscopy peaks, depicting bio-molecular interactions like non-covalent bonds, polyproline helical factors, denaturation, hydrogen bonding etc have been used to differentiate between heat treated and unheated crystallized peptide samples. The differences

between the heated and unheated supramolecular assemblies were also investigated with infrared spectroscopy.

Mass spectrometry was used to look for any possible oxidation that could have occurred along the peptide chains due to heat treatments. DLS was used to characterize morphologies, shape and size of these supramolecular assemblies affected by temperature cycles.

The effect of electrolytes (at various pH and ionic strengths) on self-associations of $(GPO)_5$ and $(GPO)_7$ peptides have been studied via TEM. Supramolecular assemblies of $(GPO)_7$ peptides in various electrolytes have also been imaged by AFM.

4.3 Characterization studies with infrared spectroscopy, from literature

4.3.1 Collagen and collagen model peptides

X-ray crystallography has proved useful to provide insights into the high-resolution structures of basic triple helical conformations and the water mediated hydration networks in collagen mimetic peptides [70, 114]. However, Infrared spectroscopy has proved very instrumental in bridging and transferring structural information from X-ray diffraction [142]. The peaks of IR spectroscopy denote the intermolecular interactions between protein molecules with different origins, such as electrostatic, hydrophobic and van der Waals forces [143].

4.3.2 Interpretation of amide peaks from literature

In vibrational spectroscopic studies, the most useful peak used for study is the amide I peak which is a C=O peptide bond stretch. It is located near 1650 cm^{-1}. From literature, amide bands at 1645–1657 cm^{-1} were specified as coils and the amide peaks at 1660 cm^{-1} were marked as triple helices [144]. Bands at 1690 cm^{-1} have been marked to be the helices of aggregated collagen like peptides [145]. From studies of collagen model polypeptides by Lazarev et al, if there were peaks from 1628 to 1633 cm^{-1}, those were considered as characteristic of the absorption of imine carbonyls and thus peak at ~ 1631 cm^{-1} symbolised hydroxyproline concentration [146]. Mello et al. have found that the band peaks at 1631 cm^{-1} represent the uncoiling of the collagen triple helix [147].

In collagen, the shift of amide I, II and III peaks to low wavenumbers and the broadening of amide I peak were associated with increased intermolecular interactions

(hydrogen bonding) in collagen [148]. Table 4.1 shows all amide bonds that are present in an IR spectrum [86]. Table 4.2 shows the IR analysis on all collagen and collagen model peptides reported until now.

Table 4.1 Significant amide peaks found in an infrared spectrum of peptides, polypeptides and proteins [86]

	Characteristic infrared bands of peptide linkage	
Designation	Approximate frequency (cm−1)	Description
Amide A	3300	NH stretching
Amide B	3100	NH stretching
Amide I	1600−1690	C=O stretching
Amide II	1480−1575	CN stretching, NH bending
Amide III	1229−1301	CN stretching, NH bending
Amide IV	625−767	OCN bending
Amide V	640−800	Out-of-plane NH bending
Amide VI	537−606	Out-of-plane C=O bending
Amide VII	200	Skeletal torsion

Table 4.2: Amide I peak frequencies of collagen and collagen related models [147]

Material characteristics	Peak frequencies (cm^{-1})	References
Collagen (Bovine cartilage) Proteoglycan	1655 1640	[149]
Collagen amide I (Porcine skin)	1654	Twardowski and Anzenbacher (1994);
Elastin (Common fibrous protein apart from collagen)	1655	Singh (2000) [150]
Collagen from Porcine pericardium tissue (tissue surrounding heart)	1654	[151]
Elastin	1655	
Heart tissue proteins adopting a predominantly α helical configuration	1650	Liu et al. (1996)
Native rat and cod skin collagens	1630 (shoulder), 1656 (central peak)	[146]
Polyproline II in D_2O solution (20 °C)	1632	
Polyproline II dried film	1640	
Polyproline II after heating	1650	
(Gly-Pro-Ala)n in D_2O solution (80 °C)	1630, 1656	
Z-$(Gly\text{-}Pro\text{-}Pro)_8$-OMe dried film	1640, 1665, 1693	
Z-$(Gly\text{-}Pro\text{-}Gly)_4$-OMe dried film	1637, 1655, 1687	
Native monomeric collagen type I at 4 °C	1650	Payne and Veis (1988)
De-convolved native collagen at 4 °C (triple helix conformation)	1660	

4 % aqueous gelatin		
4 °C	1641	
20 °C	1635	
50 °C	1631	
Collagen		
4 °C	1659	
20 °C	1659	
50 °C	1637	
Bovine pericardium collagen	1630, 1660	Sellaro et al. (2007)
Bovine pericardium collagen after changes in helicity caused by mechanically induced molecular fatigue	1632	
Collagen (Bovine derived)	~1635	[152]
LAH_4 peptide (Synthetic helical structure)	1657	Vigano et al. (2000)

Bryan and Brauner et al. reported three non-equivalent vibrations for Amide I due to the presence of the C=O groups within $(Pro\text{-}Pro\text{-}Gly)_{10}$ peptides [142]. Lazarev et al. saw these three vibrations at Amide I and assigned 1628 cm^{-1} vibration as C_GO_G and the near 1644 cm^{-1} vibrations as C_XO_X and C_YO_Y ,where G is glycine and X and Y were either prolines or hydroxyprolines [146]. Studies of Jackson and Mantsch et al., assigned the Amide I peak at 1628 cm^{-1} as the intermolecular hydrogen bonding within POG triple helix [153]. From studies of Lazarev et al., the three peaks assigned for Amide I region of polyproline were 1650 cm^{-1}, 1640 cm^{-1} and maximal absorption was seen at 1632 cm^{-1} [146].

4.4 Results and Discussion

4.4.1 IR spectroscopy on ‘unheated’ and ‘heated’ crystals of GPO collagen peptides

Heat treated crystals were prepared by heat treating the peptide solution (prepared in 10% acetic acid) which would be later crystallized. Post heat treatment, the usual crystallization procedure from Chapter 2 (conditions of crystallisation) was followed. i.e. peptide solutions

were mixed with the reservoir solutions in 1:1 ratio and the droplet was set up for crystallization by the hanging drop vapor diffusion method. The unheated crystals were prepared following the same crystallization procedures as above, at the same concentrations as discussed in (Chapter 2) but without heat treatment.

In infrared spectroscopy, for mounting the biological samples, two kinds of windows were used (1) CaF_2 windows and (2) diamond windows. For the kinds of crystal samples and nano-assembled samples used in this chapter, diamond window are best suited. The samples would sandwich perfectly focus the beam in the diamond windows and facilitate infrared spectrums, without any disturbances. Bruker V80v Fourier transform infrared (FTIR) spectrometer and Hyperion 2000 IR microscope were used at the Australian synchrotron, to record data with high signal-to-noise ratio [154].

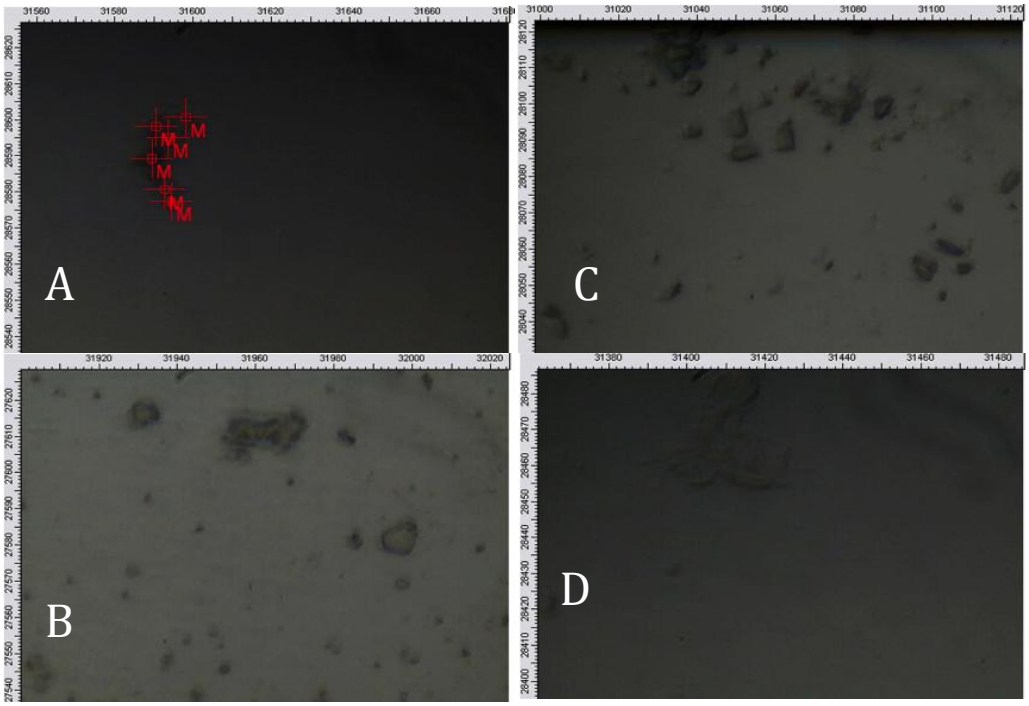


Figure 4.1 IR microscopy on unheated crystal samples sandwiched in the diamond windows; the red squares are the aperture sizes, to focus the beam (A) $(GPO)_3$ peptides, (B) $(GPO)_5$ peptides, (C) $(GPO)_7$ peptides, (D) $(GPO)_{10}$ peptides

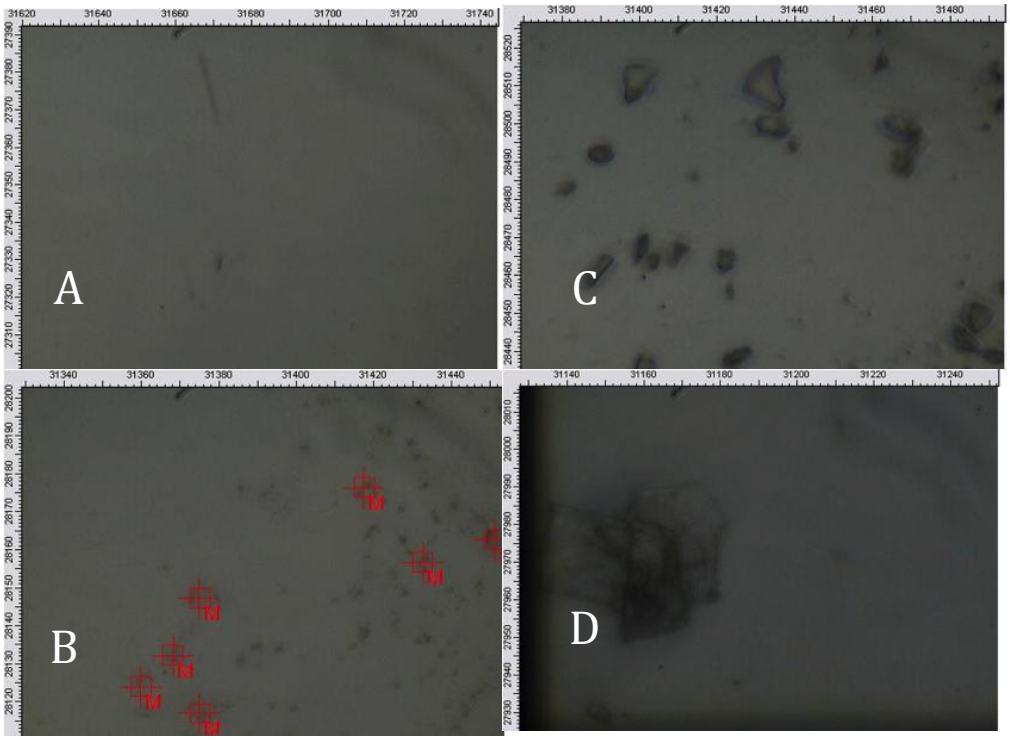


Figure 4.2 IR microscopy on heated crystal samples sandwiched in the diamond window; the red squares are the aperture sizes, to focus the beam (A) $(GPO)_3$ peptides (B) $(GPO)_5$ peptides (C) $(GPO)_7$ peptides (D) $(GPO)_{10}$ peptides

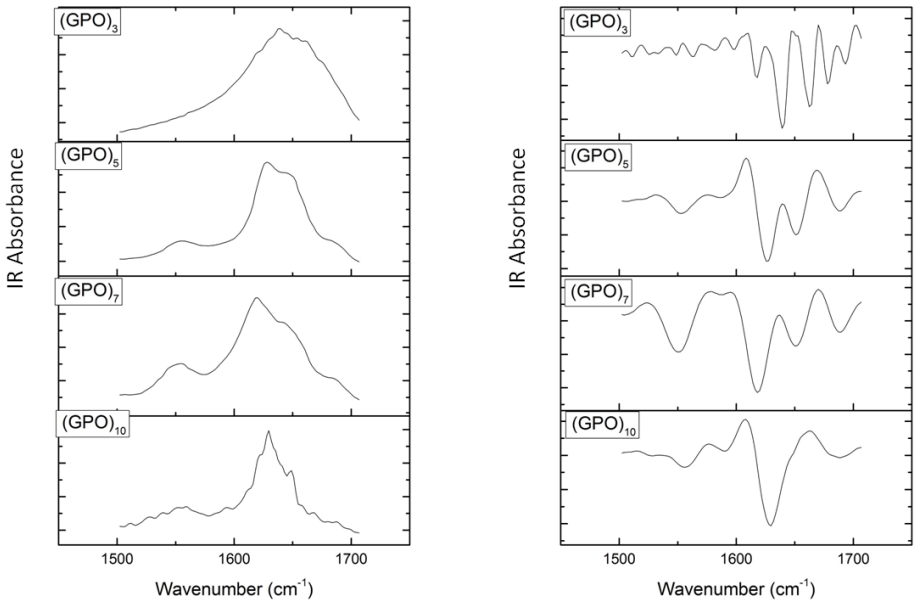
4.4.1.1 Infrared microscopy

Each microscope was equipped with a motorized sample stage and could accommodate samples up-to approximately 10×10 cm in size. A spatial resolution of ~5 microns, with an aperture (beam size) of 5 microns was used. It was observed that this setting could obtain good spectra in 10 seconds per point. Figure 4.1 and Figure 4.2 show the IR microscopy results on crystal samples. Red squares show the 5-micron aperture sizes. Figure 4.3 and Figure 4.4 are the IR spectra of crystal samples. These were generated after subtraction of the water spectra and crystal's background spectra from the raw IR spectra.

A second derivative was used to form a second derivative spectrum. Raw infrared spectra, which gets affected by noise can give slight wrong predictions for the peak maxima. However, second derivative is a mathematical method to pick the correct peak maximum (peak frequencies). Second derivatives were calculated by Bruker OPUS software, OPUS v7.2 [154].

4.4.1.2 Data interpretation

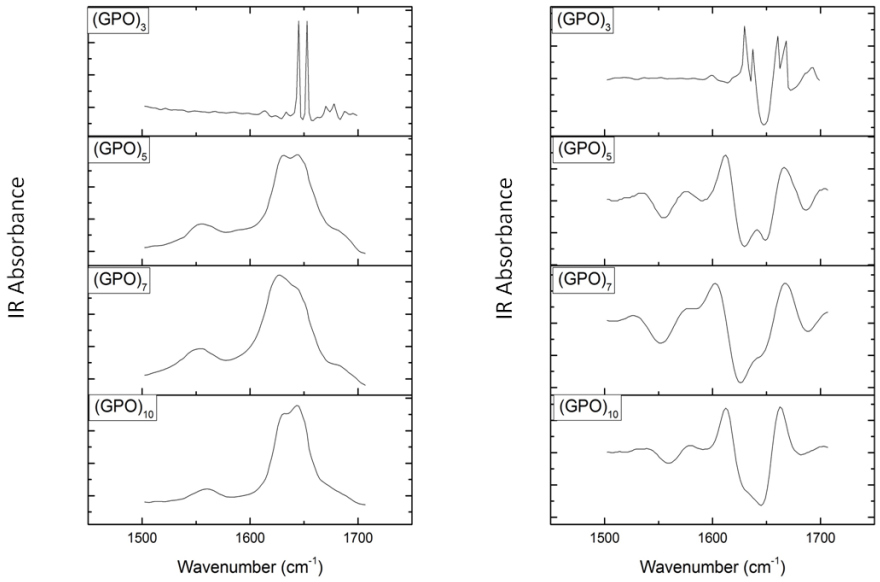
(A) Infrared spectra



A B

Figure 4.3 IR spectroscopy of un-heated crystals (A) Raw IR spectra (water subtracted)

(B) Second derivative spectra



A B

Figure 4.4 IR spectroscopy of heated crystals (A) Raw IR spectra (water subtracted) (B) Second derivative spectra

From graphs in Figure 4.3 (Table 4.3) and Figure 4.4 (Table 4.4), it was seen that for $(GPO)_3$ peptides, the amide I peak shifts from 1640 cm^{-1} in unheated samples to 1642 cm^{-1} in the heat treated samples. Since 1640 cm^{-1} and 1642 cm^{-1} peaks are associated with random coils, no absolute structure was noticed.

Table 4.3 Peaks extracted from IR spectrum and second derivative spectrum of unheated crystals

Infrared Spectroscopy			Second derivative		
Peptides	Amide I	Amide II	Peptides	Amide I	Amide II
$(GPO)_3$	1639 1651		$(GPO)_3$	1640, 1663, 1678	
$(GPO)_5$	1628	1556	$(GPO)_5$	1627, 1651 1688	1553
$(GPO)_7$	1619 (aggregation)	1554	$(GPO)_7$	1618 (aggregation) 1651	1551
$(GPO)_{10}$	1629 Shoulder-1649	1559 1594	$(GPO)_{10}$	1629 1688	1556, 1590

For unheated $(GPO)_5$ peptides, the amide I peak was seen at 1628 cm^{-1}. This is characteristic of H-bonding [147, 153]. The peak shifts to higher wavenumbers of 1631 cm-1 for the temperature treated samples and it was predicted that heat must have brought about uncoiling of $(GPO)_5$ peptides because from the studies of Campos et al. signal at 1631 cm^{-1} is associated with uncoiling of the triple helix.

For $(GPO)_7$ peptides, the amide I peak at 1618/1619 cm^{-1} shifts to 1626/1627 cm^{-1}. From literature, peaks at both 1618 cm^{-1} and 1626 cm^{-1} have been associated with macromolecular aggregation [155]. Aggregation is prominently seen in $(GPO)_7$ peptides before and after heating.

Table 4.4 Peaks picked from IR spectrum and second derivative spectrum of heated crystals

Infrared Spectroscopy			Second derivative		
Peptides	Amide I	Amide II	Peptides	Amide I	Amide II
$(GPO)_3$	1645, 1653	-	$(GPO)_3$	1642(coils) 1664 1626	
$(GPO)_5$	1631 (H-bond)	1555.6	$(GPO)_5$	1630 1687	1553
$(GPO)_7$	1627	1554	$(GPO)_7$	1626 1688	1551
$(GPO)_{10}$	1644 1631	1560	$(GPO)_{10}$	1645 1682	1559

Table 4.5 IR spectroscopy; Amide I peak shifts in collagen model peptide crystals, after heat treatment. Peptides with significant differences have been shown with *.

Peptides	Unheated crystals (cm^{-1})	Heated crystals (cm^{-1})
$(GPO)_3$	1640	1642
$(GPO)_5$	1627	1630
$(GPO)_{7*}$	1618	1626
$(GPO)_{10*}$	1629	1645

For $(GPO)_{10}$ peptides, unheated base peak was visible at 1629 cm^{-1} (following the second derivative). Tiny shoulders were seen at 1648 and 1687 cm^{-1}. All three vibrations of poly-proline spectra were clearly located. The amide I peak seen at 1629 cm^{-1} which is the characteristic of H-bonding got shifted to 1644/1645 cm^{-1} in the heat-treated samples. Peak at 1645 cm^{-1} are denoted as random coils [156]. Hence, uncoiling brought about by heating is prominent in $(GPO)_{10}$ peptides. For all peptides, the wavenumber shift from lower to higher wavenumbers, after heat treatments is recorded in Table 4.5.

4.4.2 Effect of temperature cycles on supramolecular self-assemblies

Aggregation prone zone is a temperature zone in which the loosened intermediate structure further associates to form higher order structure (supramolecular assembly).

This phenomenon has been reported in literature via turbidity studies, which usually occurs just below melting point [48, 60]. From the studies of Kar et al, the best possible supramolecular assemblies have been reported for $(GPO)_{10}$ peptides at 70 °C after incubation for 8 minutes and then cooling down to room temperature, at pH 7 and 20 mM PBS buffer [48].

$(GPO)_3$ $(GPO)_5$ $(GPO)_7$ $(GPO)_{10}$

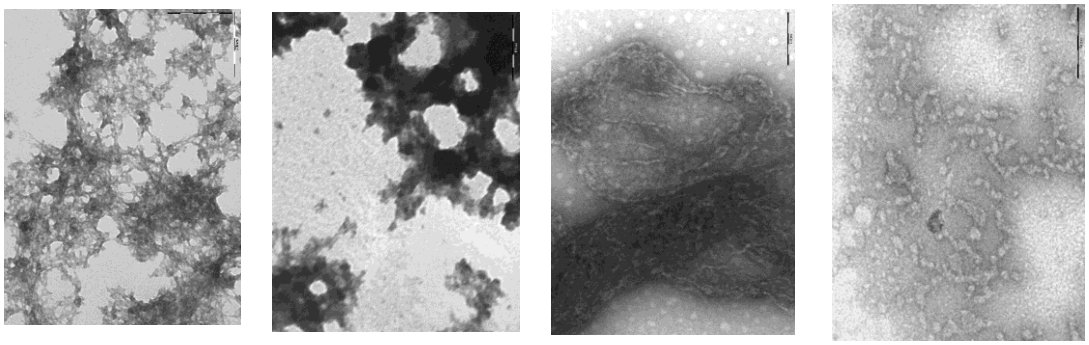


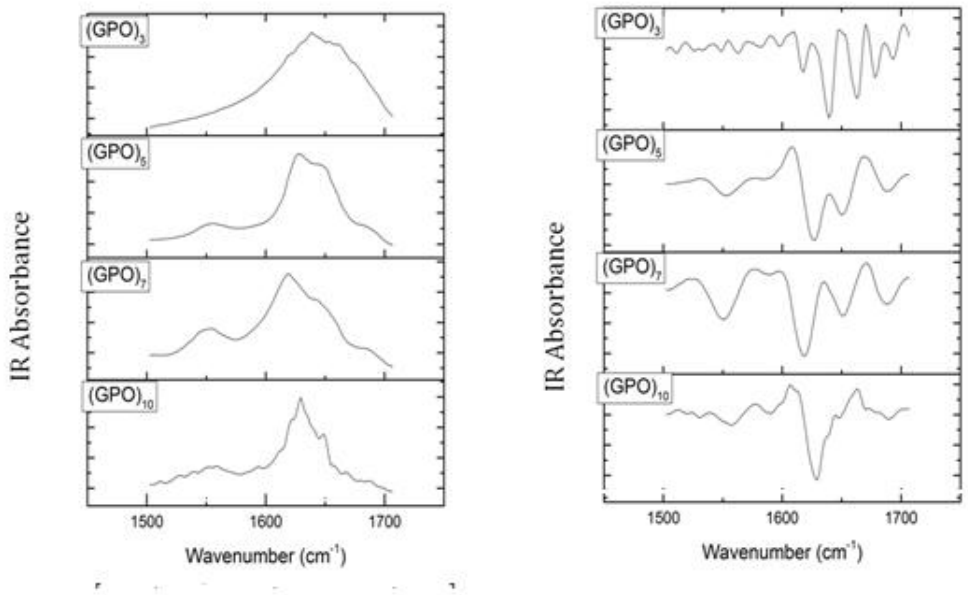
Figure 4.5 TEM analysis; supramolecular assembly formations for $(GPO)_3$, $(GPO)_5$, $(GPO)_7$, $(GPO)_{10}$ peptides, heat treated at 70 °C, incubated for 8 minutes and cooled down. 20 mM PBS buffer was used at pH 7. $(GPO)_3$ $(GPO)_5$ peptides show random aggregates, while $(GPO)_7$ shows significant supramolecular assemblies followed by $(GPO)_{10}$ peptides. TEM artefacts were differentiated from supramolecular assemblies based on following references [131], [132].

Similar experimental conditions were followed to obtain supramolecular assemblies. Incubation temperatures were selected, based on the melting temperatures of $(GPO)_3$, $(GPO)_5$, $(GPO)_7$ and $(GPO)_{10}$ peptides i.e. 37 °C, 50, °C 60 °C and 70 °C. At a concentration of 7 mg/ml in 20 mM PBS (pH 7), samples were incubated in their respective temperatures for 8 minutes and cooled down to room temperature (20 °C) to visualise supramolecular assemblies by TEM. Detailed analysis and description of these temperature treatments have been further discussed in section 4.4.2.2, where the longer chain length of peptides, $(GPO)_{10}$ and $(GPO)_7$ were selected for further experimentations.

Significant supramolecular assemblies were recorded for both $(GPO)_7$ and $(GPO)_{10}$ peptides after temperature treatments at 70 °C. $(GPO)_5$ and $(GPO)_3$ peptides were also subjected to same conditions (see Figure 4.5). However, supramolecular assembly formations were not seen for $(GPO)_5$ and $(GPO)_3$ peptides, rather these appeared to adopt random

aggregate formations. Infrared spectroscopic analysis was quickly performed on all the resulting GPO structural assemblies, to characterize the changes arising in the biomolecular interactions due to the temperature treatments.

4.4.2.1 Analysis of supramolecular self-assemblies by IR spectroscopy



A B

Figure 4.6 Reference unheated supramolecular assemblies (A) Infrared spectra (B) Second derivative spectra

Infrared spectroscopy was performed on the supramolecular assemblies from Figure 4.5, for $(GPO)_3$, $(GPO)_5$, $(GPO)_7$, $(GPO)_{10}$ peptides (see Figure 4.7A and Figure 4.7B, Table 4.7). Samples were prepared at 7 mg/ml at 70 °C incubation for 8 minutes in 20 mM PBS buffer. The reference-unheated structures are shown in Figure 4.6A and Figure 4.6B, Table 4.6. These unheated samples were prepared at 7 mg/ml concentrations and were incubated at room temperature for 8 minutes, in 20 mM PBS buffer.

The spectra from the supramolecular assemblies of (GPO) peptides have been recorded in (Table 4.7) and the amide peaks for the unheated and heated supramolecular assemblies are shown. Amide I peaks from the reference samples, the unheated supramolecular assemblies (i.e. prepared at 7 mg/ml concentrations and incubated at room temperature for 8 minutes, in 20 mM PBS buffer) are shown at Table 4.6.

For $(GPO)_3$ triple helices, shift in the absorption was seen for the amide I band from 1627 cm^{-1} (depicting H-bonds), for unheated assemblies to 1641 cm^{-1} (depicting random coils) for heated assemblies. For $(GPO)_5$ peptides, amide I was seen at 1627 cm^{-1} which also represents the H-bonding in triple helices. After heating the amide I peak shifts to 1644 cm^{-1}, which also represents the random coils at ~ 1640s. Perhaps, the breakage of H-bonds and random coil formation was caused by heating.

Table 4.6 IR spectroscopy on unheated GPO supramolecular assemblies

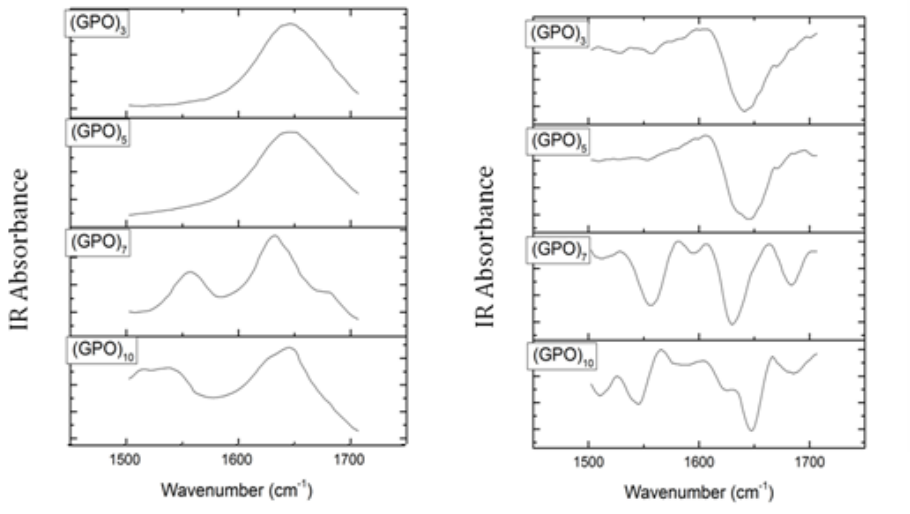
Infrared Spectroscopy		
Peptides	Amide I	Amide II
$(GPO)_3$	1646	
$(GPO)_5$	1656	
$(GPO)_7$	1633	1557
$(GPO)_{10}$	1645	1538, 1516
Collagen	1653	

Second derivative		
Peptides	Amide I	Amide II
$(GPO)_3$	1641	
$(GPO)_5$	1644 1686	1553
$(GPO)_7$	1630 1684	1553 1511
$(GPO)_{10}$	1648 1626	1545 1510
Collagen	1651	1553

Table 4.7 IR spectroscopy on heated GPO supramolecular assemblies

Infrared Spectroscopy		
Peptides	Amide I	Amide II
$(GPO)_5$	1628	1556
$(GPO)_7$	1619 1688	1554
$(GPO)_{10}$	1629 1648.6	1559 1593.9
Collagen	1647	1561

Second derivative		
Peptides	Amide I	Amide II
$(GPO)_5$	1627 1650 1619	1553
$(GPO)_7$	1651	1551
$(GPO)_{10}$	1629 1647	1557, 1590
Collagen	`1656	1598



A B

Figure 4.7 Heated supramolecular assemblies (A) Infrared spectra (B) Second derivative spectra

A large shift of wavenumber from 1619 cm^{-1} to 1630 cm^{-1} was observed for the $(GPO)_7$ peptides. This absorbance suggests macromolecular aggregation. These results are in accordance with the IR results for unheated and heated $(GPO)_7$ crystals. For $(GPO)_{10}$ peptides, a large shift of wavenumber was seen to occur for the band at 1629 cm^{-1} for unheated samples (depicting presence of H-bonds) to 1648 cm^{-1} for heated samples (depicting random coils). Thus indicating breakage of intermolecular hydrogen bonding upon heating. Please see Table 4.8 for details.

Table 4.8 Studying Amide I peak shifts in unheated and heated treated (GPO)n supramolecular assemblies and collagen; by IR spectroscopy

Peptides	Amide I for Unheated structures (cm-1)	Amide I for Heated structures (cm-1)
$(GPO)_3$	-	1641
$(GPO)_5$	1626	1644
$(GPO)_7$	1619	1630
$(GPO)_{10}$	1629	1648
Collagen	1656	1651

4.4.2.2 Effect of temperature cycles on the supramolecular assemblies of $(GPO)_{10}$ and $(GPO)_7$ peptides

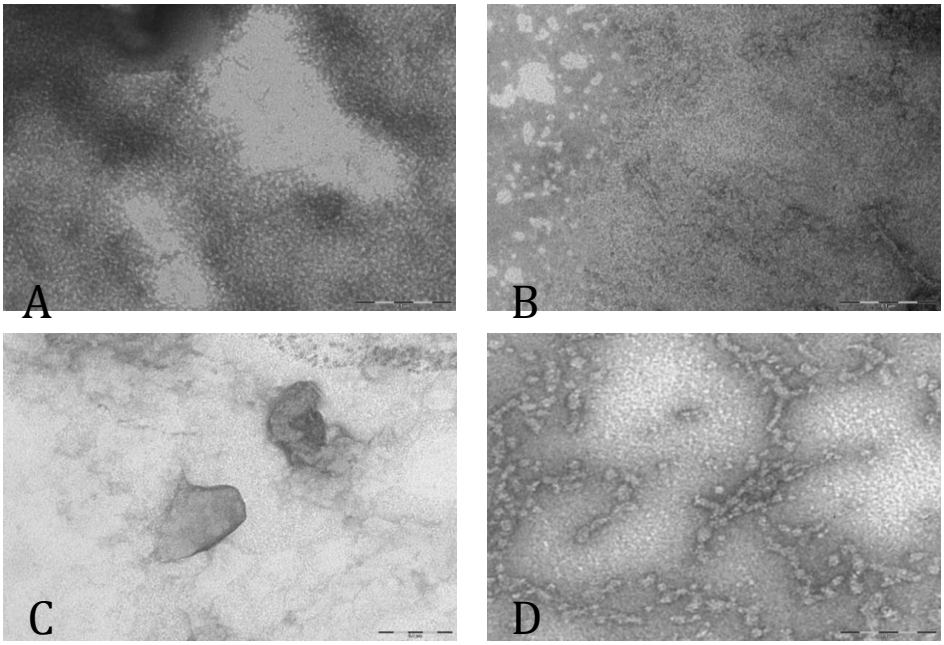


Figure 4.8 $(GPO)_{10}$ supramolecular assembly (A) Aggregations were seen at 37 °C (B) No clear assemblies at 50 °C (C) No significant structure was seen for 60 °C (D) Higher order structures were observed for 70 °C. (These have been replicated). TEM artefacts were differentiated from supramolecular assemblies based on following references [131], [132].

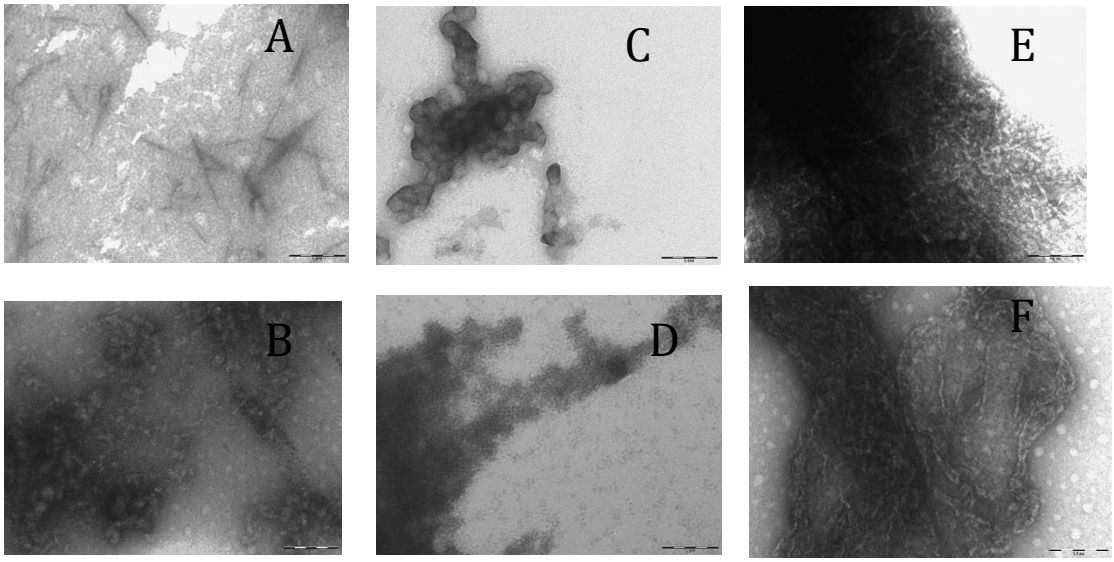


Figure 4.9 Supramolecular assembly of $(GPO)_7$ peptides (A) Nano-rods observed at 37 °C (B) branched filamentous structures at 37 °C (C) no clear assembly was seen at 50 °C (D) no clear assembly was observed at 60 °C (E) Fibrillar assembly at 70 °C (F) ordered supra-structures at 70 °C. (These have been replicated).

Structural changes due to the temperature cycles, at 37 °C, 50 °C, 60 °C and 70 °C have been discussed in this section.

Supramolecular assemblies of $(GPO)_{10}$ peptides (made by temperature cycles) have been shown in Figure 4.8 A, B, C and D. Ordered assemblies were seen at 70 °C, which have already been analysed by IR spectroscopy in the previous sections. Aggregations were also seen at 37 °C and no clear assemblies were seen at 50 °C or at 60 °C.

In Figure 4.9A, B, C and D, supramolecular assemblies of $(GPO)_7$ peptides are shown. Nano-rod structures were observed for the peptides at 37 °C. As far as it is known, these structure have not been reported in literature. Since this was the near melting temperature for $(GPO)_7$ triple helices, the hypothesis established in the previous chapters and the research predictions by Kar et al. has been reproved, that is, higher order structure formations due to self-associations occur at a temperatures just below the melting point of triple helices [48]. Sufficient similarities have been found between these newly discovered $(GPO)_7$ nano-rods and the nano-rods formed by collagen peptides lacking hydroxyproline, from studies of Krishna and Kiick et al. [60].

No clear assemblies for $(GPO)_7$ peptides were seen at 50 °C and 60 °C; but supramolecular assemblies for these peptides were observed when heat treated at 70 °C. These could have happened because of the possible self-association of the unwound monomers at 70 °C. These have already been analysed by IR spectrocopy in the previous section (Figure 4.9F).

4.4.2.3 Effect of temperature cycles on the supramolecular assemblies of amidated $(GPO)_7$ peptides

Random aggregates were noticed for the heat treated peptides (after incubation at 70 °C, for 8 minutes and cooling down to room temperature, 20 °C (Figure 4.10B); whereas no supramolecular assembly formations were seen for unheated amidated $(GPO)_7$ peptides (Figure 4.10A).

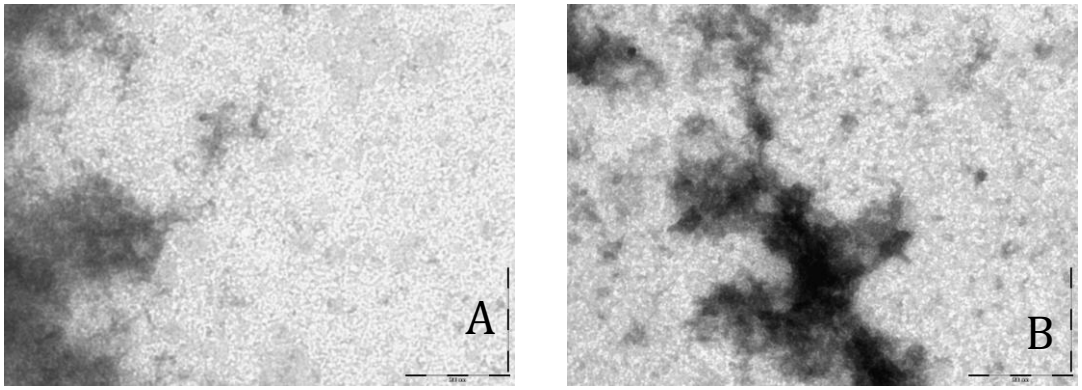


Figure 4.10 (A) Supramolecular assembly formations for amidated $(GPO)_7$ peptides at 7 mg/ml in 20 mM PBS buffer, pH 7(A) unheated amidated $(GPO)_7$ showed no supramolecular assemblies (B) heat treated at 70 °C, incubated for 8 minutes and cooled down; aggregates were noticed.

4.4.3 Effect of heat treatments on the association kinetics of $(GPO)_7$ and $(GPO)_{10}$ supramolecular assemblies; Analysis by DLS

DLS analysis was performed on the temperature-favored supramolecular assemblies at 7 mg/ml which were incubated at 70 °C for 8 minutes in 20 mM PBS (pH 7). These assemblies have been reproduced from section 4.4.2.1. Reference set of samples were also prepared at 7 mg/ml at 8 minutes of incubation time at room temperature (20 °C), to study the effects of heating on the association kinetics of assemblies.

4.4.3.1 Helices

Kar et al have shown that $(GPO)_{10}$ peptides show the presence of single population of species with R_h (hydrodynamic radius) of ~ 2 nm at 37 °C at 60 minutes of incubation time, these were the single population of native triple helices [48].

In Figure 4.11 and Table 4.9, possible helices of ~ 1-2 nm for the $(GPO)_7$ peptides were seen, when incubated at room temperature. These were also visible for the heat-treated samples. For the unheated $(GPO)_{10}$ peptides, possible helices or initial helical aggregations were also noticed. These fused structures were of ~ 4 nm size. In the heat-treated $(GPO)_{10}$ samples, the size of these species had reduced to ~ 2.3 nm.

4.4.3.2 Supramolecular assemblies

The reference sample for $(GPO)_7$ peptides, has second aggregate population of size ~ 44.5-55.4 nm and 190 nm. After heating, the 190 nm size of species had reduced to supramolecular assemblies of ~ 139 nm.

From previous studies of Kar et al., formation of smaller associated structures prior to the formation of larger aggregates have been noticed [48]; and the presence of higher molecular weight DLS peak indicates the presence of second population of R_h values of ~ 65-70 nm for $(GPO)_{10}$ peptides [48]. Kar and Amin et al. have reported triple helices of hydrodynamic radius ~1.9 nm at 4 °C at 7 mg/ml in pH 7, PBS buffer for $(GPO)_{10}$ peptides but no aggregations were seen when their system was heated till 58 °C [48].

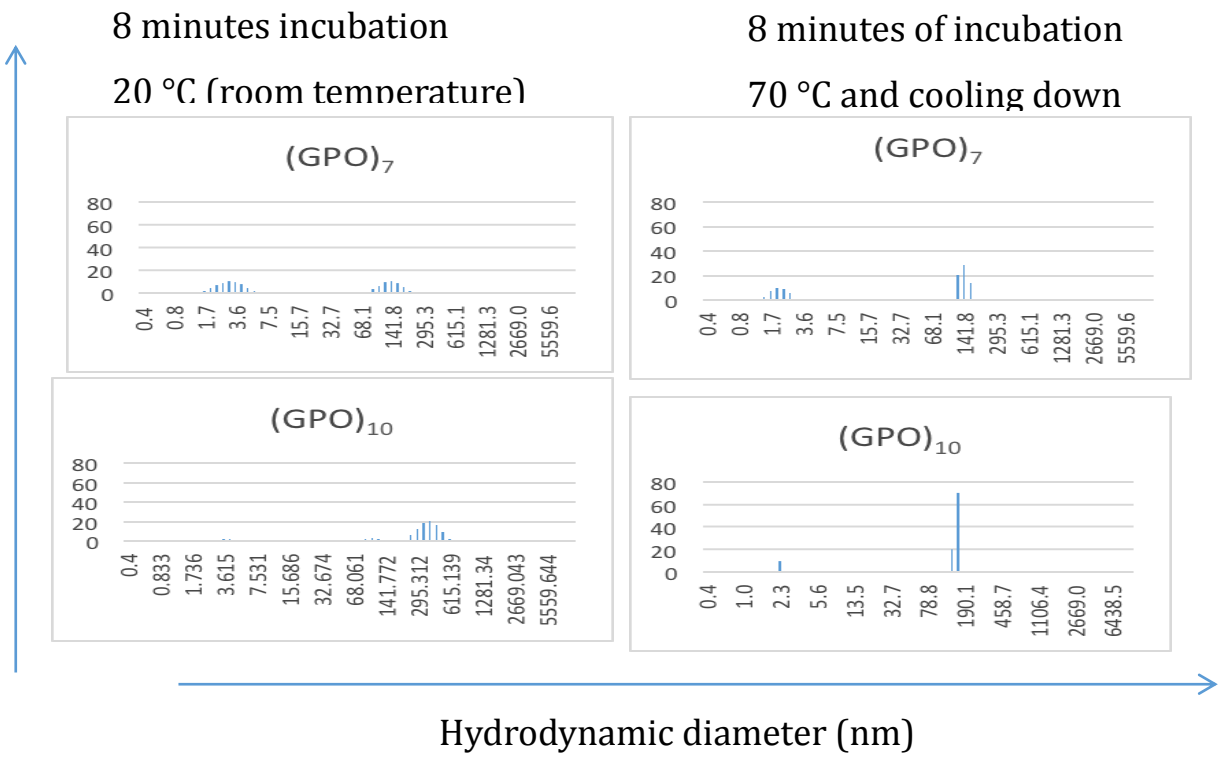


Figure 4.11 DLS (intensity vs hydrodynamic diameter) on $(GPO)_7$ and $(GPO)_{10}$ peptides at 7 mg/ml in 20 mM PBS buffer at pH-7; Left lane are the samples, incubated at 20 °C (room temperature) for 8 minutes. The right lane are the samples incubated at 70 °C for 8 minutes and cooled down.

For the $(GPO)_{10}$ peptides, the associated supramolecular assemblies were of size ~390 nm diameter. This large size was possibly due to the association of smaller species. However, after heat treatment, the supramolecular assemblies of $(GPO)_{10}$ peptides had reduced to ~ 159 nm size.

Particle size analysis via DLS was performed on a collagen mimetic peptide specifically designed to form microfibrils, homotrimer $(GPP)_3GPRGEKGERGPR(GPP)_3$ GPCCG,

(where C is cystine, R is Arginine and E is Glutamic acid). The mixture formed three species of diameters 5.2 nm, 226 nm and 5.6 um.

$(GPO)_{10}$ and $(GPO)_7$ peptides formed triple helices at room temperature, along with aggregations. The significant finding from this experiment is that after heat treatments, decrement of associated species was repeatedly noticed for the peptides. It is speculated that (from the IR spectroscopic experiments section 4.4.2.1 on the same samples), reduction of these supramolecular assemblies could have resulted from the breakage of intra and inter helical H-bonds, after heat treatment. This size reduction of assemblies has also been backed up with the AFM imaging results on $(GPO)_7$ supramolecular assemblies, which have shown size reduction after heat treatments in section 4.4.5.6.

Table 4.9 Self-associated $(GPO)_7$ and $(GPO)_{10}$ peptides. Measurement by DLS (Diameter in nm). Significant decrease in size of species, upon heating is shown by *.

	8 minutes incubation at room temp (nm)	8 minutes incubation at 70 ^{0}C and cooling down (nm)
	$(GPO)_7$	$(GPO)_7$
Peaks	1.1 44.5, 54.4 190*	1.1 - 139*
	$(GPO)_{10}$	$(GPO)_{10}$
Peaks	4.005 103.7 390.3*	2.314 - 159.2*

4.4.4 Small angle X-ray scattering (SAXS) analysis on GPO supra-assemblies

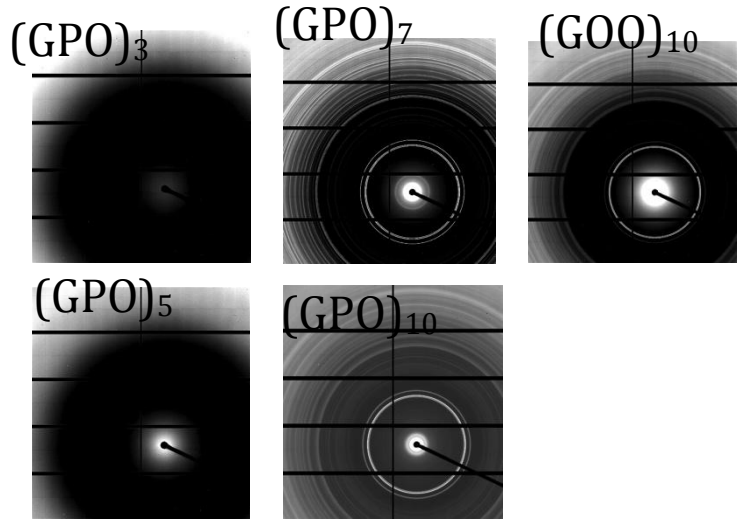


Figure 4.12 Micro-crystals were seen for $(GPO)_7$, $(GPO)_{10}$ and $(GOO)_{10}$ while $(GPO)_3$ and $(GPO)_5$ were soluble.

There have been several reports on collagen like peptides that self-associate into higher order structures including $(POG)_{10}$ [110]; some self-associate into liquid crystalline arrays of triple helical peptides [48, 111]. To examine the presence of these crystalline arrays with (GPO) peptides within this research, small angle X-ray scattering was performed (as shown in Figure 4.12).

Gurkan et al. have studied the two-dimensional SAXS patterns for natural collagen and tendon and analysed their D banding periods [157]. A fibril is formed due to staggered arrangement of many collagen molecules. In this arrangement, after a period a next collagen molecule overlaps, which forms a gap before the next collagen molecule, these are known as D-spacings. D-spacing is a significant characteristic in determining the structure and function of a protein or peptide self-assembly. From previous studies, collagen and native tendon's D-spacings or the overlap spacings have been found to be ~ 61-62 nm (based on $d=2\pi/q$) [157].

No two-dimensional SAXS analysis has been performed on collagen mimetic peptides to study the D banding periods until date. This might be the first initiative to study the (GPO) self-assemblies. The reason behind scarcity of literature in this field could be that, not many collagen mimetic peptides have successfully formed ordered fibrils.

The scattering experiments in this research were performed at the Australian Synchrotron (Melbourne), in collaboration with Dr Celine Valery. The employed photon energy of X-rays was 12 keV (λ = 1.03 Å). Two sample to detector distances were chosen to measure both finer molecular details and the aggregates. Samples were loaded in a quartz capillary. Scattering profiles were carried out using ScatterBrain Analysis v1.0.3 software, developed at EMBL, Hamburg.

4.4.4.1 Sample preparation

Samples were prepared in concentrations of 50 to 300 mg/ml in the following buffers; 150 mM NaCl, 10 % acetic acid, 20 mM PBS buffer and 100 mM PBS buffer. The 2 mm capillaries were centrifuged at 3500 rpm. White gels were seen in the capillaries for $(GPO)_7$ peptides at 300 mg/ml, in 100 mM PBS buffer, for $(GOO)_{10}$ peptides at 300 mg/ml, in 100 mM PBS buffer and for $(GPO)_{10}$ peptides at 200 mg/ml, in 100 PBS buffer.

Thick transparent gels were see in capillaries with all the $(GOO)_{10}$ peptide samples in 100 mM PBS buffer, 10 % acetic acid buffer and in 150 mM NaCl buffer at 200 mg/ml concentrations. These were also seen for $(GOO)_{10}$ peptides in 150 mM NaCl and in 10 % acetic acid, at concentrations of 300 mg/ml.

Table 4.10 SAXS Analysis on (GPO) peptides

Peptides	Concentrations in 100 mM PBS buffer	Q (A^{-1})	Q (1/nm)	Intensity	Distance (Å)
$(GPO)_{10}$	200 mg/ml	0.512	5.12	11	1.22
		0.528	5.28	23	1.18
		0.58	5.8	8	1.08
$(GPO)_7$	300 mg/ml	0.512	5.12	18	1.22
		0.528	5.28	24	1.18
		0.58	5.8	12.5	1.08
$(GPO)_5$	300 mg/ml	No peak	-	-	
$(GPO)_3$	300 mg/ml	No peak	-	-	
$(GOO)_{10}$	300 mg/ml	0.508	5.08	12	1.23
		0.5	5	15	1.25
		0.572	5.72	7	1.09

Transparent gels were noticed for samples of $(GPO)_7$ peptides in 100 mM PBS buffer, $(GPO)_7$ peptides in 10 % acetic acid buffer, $(GPO)_{10}$ peptides in 10 % acetic acid at 200 mg/ml concentrations. The gels were also noticed for $(GPO)_7$ peptides in 10 % acetic acid at 300 mg/ml, for $(GPO)_7$ peptides in 150 mM NaCl buffer at 300 mg/ml and for $(GPO)_5$ peptides in 100 mM PBS buffer, at 300 mg/ml concentrations. No significant formations were seen in the following samples, $(GPO)_3$ peptides in 150 mM NaCl, $(GPO)_3$ peptides in 100 mM PBS, $(GPO)_3$ peptides in 10 % acetic acid, $(GPO)_5$ peptides in 10 % acetic acid, $(GPO)_5$ peptides in 100 mM PBS and $(GPO)_7$ peptides in 150 mM NaCl at 200 mg/ml concentrations.

4.4.4.2 SAXS Analysis

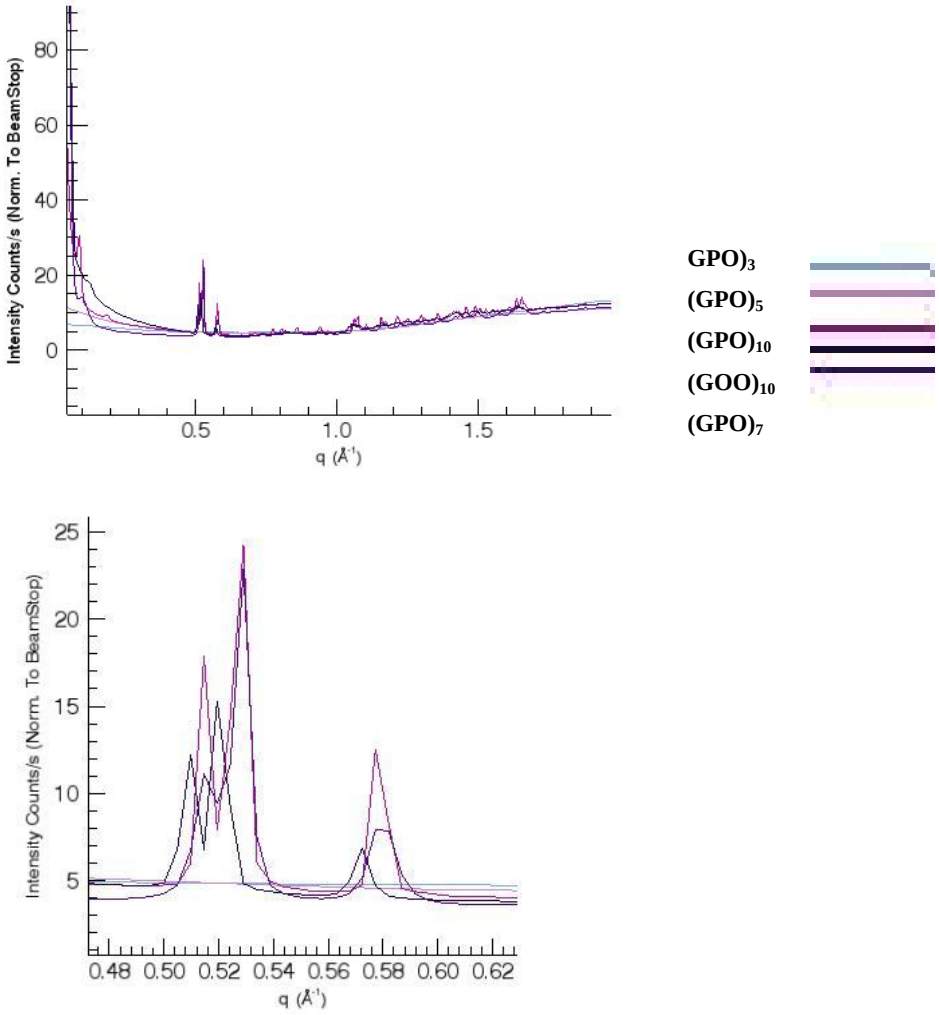


Figure 4.13 SAXS analysis on (GPO) peptides, Intensity counts vs $Å^{-1}$; (A) Diffraction pattern from 0.08 to 2.0 $Å^{-1}$ (B) Prominent Bragg peaks from the graph: three strong diffraction peaks of $(GPO)_{10}$, $(GPO)_7$ and $(GOO)_{10}$ were noticed, apart from the many random diffraction peaks because of liquid crystals.

The diffraction pattern has been recorded for reciprocal spacing q ($Å^{-1}$), from 0.08 to 2.0 $Å^{-1}$; see Figure 4.13. The range of repetitive distances d= 2π/q was from 78.5 Å to 3.14 Å. All samples exhibited powder diffraction and scattering intensities as a function of the radial wave vector, q = 4π sin θ/λ. Most clear Braggs peaks were obtained for the samples as shown in the Table 4.10 (Radial integrations have been summarised in this table).

The D-spacings for both $(GPO)_{10}$ and $(GPO)_7$ peptides were found to be 1.22 nm, whereas the D-spacing for $(GOO)_{10}$ peptides, was slightly different and was found to be ~ 1.23 nm. In previous studies on spacing, Chmielewski et al. have reported that the Hbyp peptide is composed of repeating (POG) triads modified with two bipyridine moieties possessed centre to centre spacing of 3.9 and 2.8 nm respectively, suggesting a cubic assembly of those Hbyp3 disks. This is consistent with another model of cubic packed collagen peptide triple helices with interdigitating bipyridine units [158].

4.4.5 Effect of electrolytes on $(GPO)_{5,}$ $(GPO)_7$ and $(GPO)_{10}$ supramolecular assemblies, A comparative study with collagen;

Electrolytes at various pH's and ionic strengths of 150 mM NaCl, Phosphate buffer at pH 7, 20 mM PBS buffer at pH 7, 10% acetic acid buffer at pH 3, water etc were selected. Samples were incubated for 8 minutes at room temperature (20 °C) at 7 mg/ml in their respective buffers. These specific buffers were selected because, from previous chapters it was observed that the neutral buffers like PBS, phosphate, water have been supportive of assembly formations. The high ionic strength buffers like 150 mM NaCl have been shown to enhance the assembly formations even more. The acidic acetic acid buffer was seen to support both the crystallization of all peptides as well as their assembly formations.

Muller et al. has studied self-assembly of collagen into microribbons, which showed critical dependence on the effect of pH and electrolytes [159]. The effects of buffer conditions on the in vitro self-assembly of collagen have also been noted before [160] [159]. Collagen type 1 has been assembled in various electrolyte solutions and structurally investigated by electron microscopy [161-163] [159] and also by AFM [159, 164, 165].

TEM and AFM have been used to study the changes in morphologies and the changes in the sizes of GPO supramolecular assemblies, after being exposed to various conditions of electrolytes and heat treatments (incubation for 8 minutes at 70 °C) in this chapter.

4.4.5.1 Analysis of collagen by TEM

Type 1 collagen from bovine achilles tendon has been used in this experiment. From Figure 4.14, it was seen that the change in salt concentration affects the collagen fibrillar arrangement. With phosphate buffer, the collagen exhibited fibril formation (see Figure 4.14C) and a few striations were noticed in Figure 4.14D, at a scale of 200 nm. No fibril formation was seen with acetic acid buffer (Figure 4.14A and Figure 4.14B), whereas with water, collagen fibrils were seen (Figure 4.14E and Figure 4.14F).

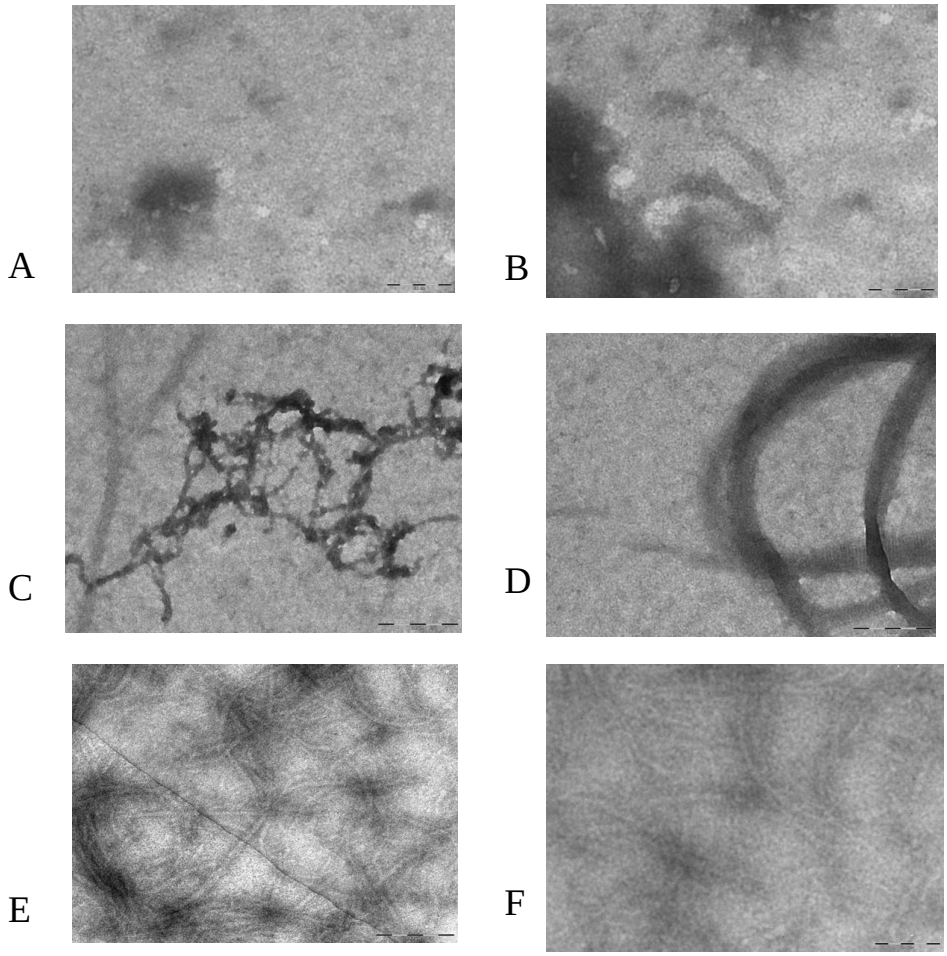
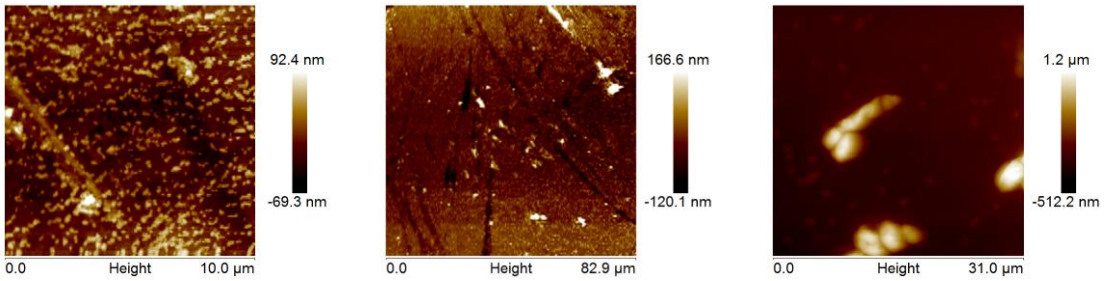


Figure 4.14 Collagen type I samples in various buffers at 7 mg/ml, pH 7 at 8 minutes of incubation time at 70 °C and cooling down to 20 °C (a) 10 % acetic acid at a scale was 100 nm (b) 10 % acetic acid at a scale of 100 nm (c) phosphate at a scale of 200 nm; fibrils were seen (d) phosphate buffer at a scale of 200 nm; noticeable striations were recorded because of staggered assembly (e) water at a scale of 100 nm; fibrils were recorded (f) water at 100 nm; fibrils were noticed.

4.4.5.2 Analysis of collagen by AFM

From previous literature, the studies of Wilson et al. were looked at, to understand AFM imaging on collagen fibrils initially [166].

From Figure 4.15 C, at a scan size of 31 µm, collagen has been analyzed and fibrils 1.2-µm high were seen. In Figure 4.15B, collagen in PBS buffer at a scan size of 83 µm, fibrillar assemblies of height ~ 166 nm were observed. In Figure 4.15A, in 20 mM PBS after 8 minutes of incubation at 70 °C and cooling down to 20 °C, collagen assemblies were observed, which were measured to be ~ 92 nm high, at a scan size of 10 µm.



A B C

Figure 4.15 AFM on collagen type I (A) Assemblies in 20 mM PBS after 8 minutes of incubation at 70 °C and cooling down to 20 °C at a scan size of 10 µm (B) Assemblies in 20 mM PBS after 8 minutes of incubation at 70 °C and cooling down to 20 °C at a scan size of ~ 83 µm; fibrillar assemblies of height ~ 166 nm were seen (C) Assemblies in water after 8 minutes of incubation at 70 °C and cooling down to 20 °C at scan size of 35 µm. AFM artefacts were differentiated from supramolecular assemblies based on following reference [132].

4.4.5.3 Analysis of $(GPO)_5$ peptides by TEM

In Figure 4.16, unheated TEM aggregates of $(GPO)_5$ peptides are shown. Buffer used for the sample peptide solutions is 150 mM NaCl in Figure 4.16A, Phosphate buffer in Figure 4.16B, water in Figure 4.16C and 10 % acetic acid buffer in Figure 4.16D. Results were inconclusive on any supramolecular assembly formations for $(GPO)_5$ peptides.

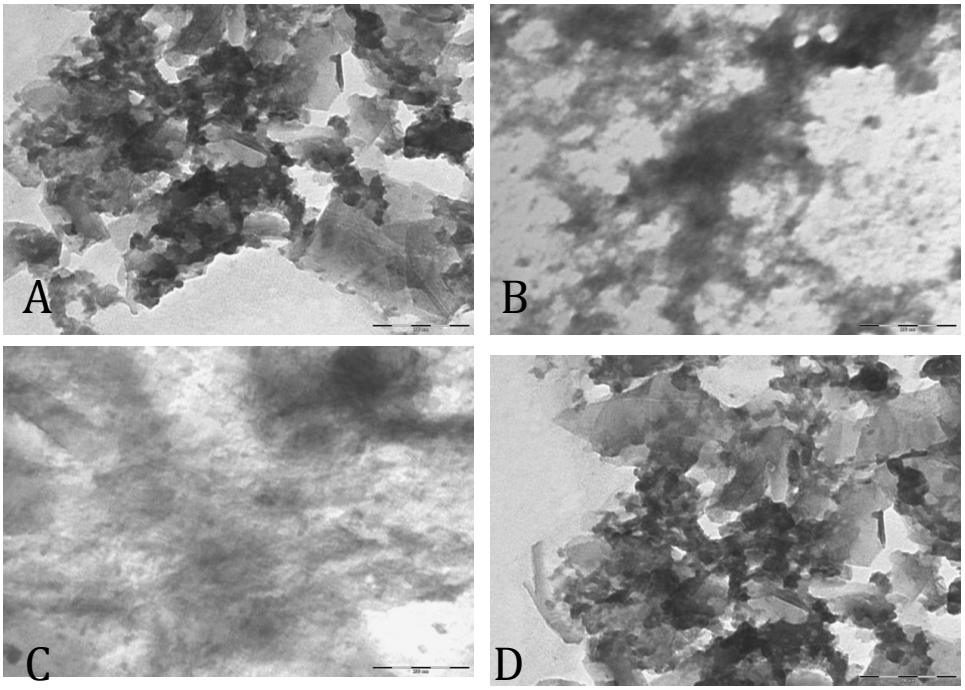


Figure 4.16 Electron microscopic images were non-conclusive for $(GPO)_5$ peptides at 7 mg/ml, after 8 minutes of incubation at 70 °C and cooling down to 20 °C (A) 150 mM NaCl buffer (B) Phosphate buffer (C) Water (D) 10 % acetic acid buffer. TEM artefacts were differentiated from supramolecular assemblies based on following references [131], [132].

4.4.5.4 Analysis of $(GPO)_5$ peptides by AFM

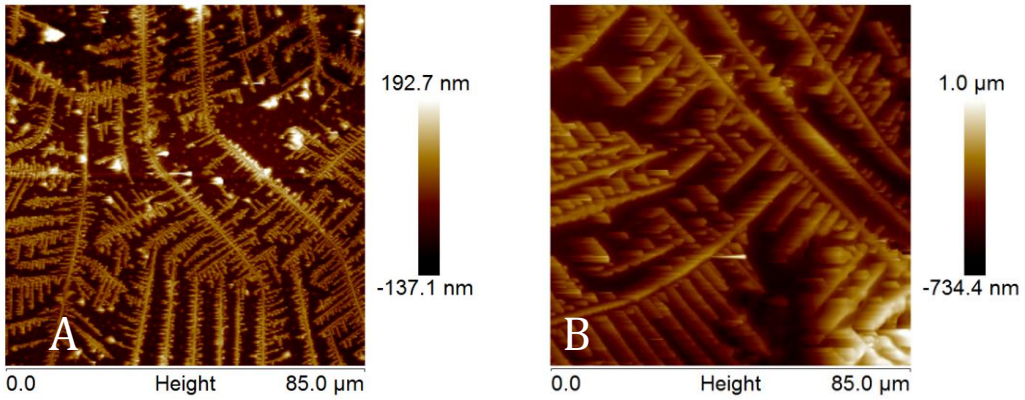


Figure 4.17 AFM images for $(GPO)_5$ peptides at 7 mg/ml in 20 mM PBS buffer. Artefacts, occurred by surface adsorption were noticed (A) $(GPO)_5$ peptides after 8 minutes of incubation at room temperature, at scan size of 85 µm (B) $(GPO)_5$ peptides after 8 minutes of incubation at 70 °C and cooling down until room temperature (20 °C) at a scan size of 85 µm. AFM artefacts were differentiated from supramolecular assemblies based on following reference [132].

Analysis by AFM was performed to study the change in morphologies and the change in the sizes of supramolecular assemblies, after heat treatment (incubation for 8 minutes at 70 °C and cooling down to 20 °C).

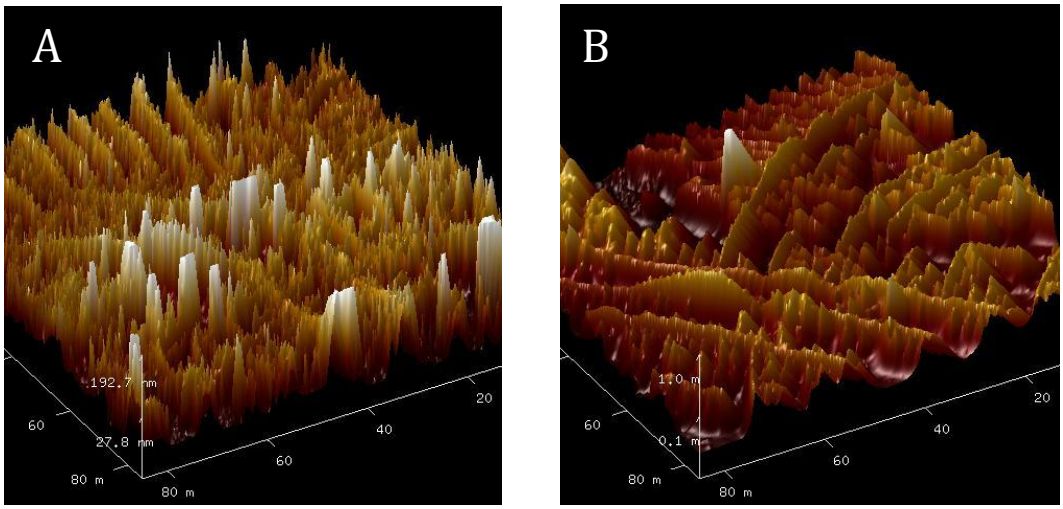


Figure 4.18 AFM height estimation of $(GPO)_5$ peptides from Figure 4.17; software used for measurements was NanoScope analysis (A) $(GPO)_5$ peptides after 8 minutes of incubation at room temperature, at a scan size of 85 µm; the assembly height was found to be ~ 190 nm (B) after 8 minutes of incubation at 70 °C and cooling down until room temperature, scan size was 85 µm. Assembly height ~ 1 µm.

Height measurements of $(GPO)_5$ peptides in 20 mM PBS high salt buffer at 7 mg/ml and scan size of 85 µm are shown in Figure 4.18. The formations look like artifacts due to surface adsorption [167]. From studies of Landoulsi et al., the artefacts have been recognised and differentiated in this research [166, 168].

4.4.5.5 Analysis of $(GPO)_7$ peptides by TEM

From Figure 4.19, aggregations were looked for , with unheated $(GPO)_7$ peptides. Peptides in 150 mM NaCl are shown in Figure 4.19A. Peptides in phosphate buffer are shown in Figure 4.19C, peptides in 20 mM PBS buffer are shown in Figure 4.19D and peptides in water are shown in Figure 4.19E. All images were non-conclusive.

Samples were further heat-treated in 150 mM NaCl (shown in Figure 4.19B) and in water (see Figure 4.19F). The images are still inconclusive of any assembly formations.

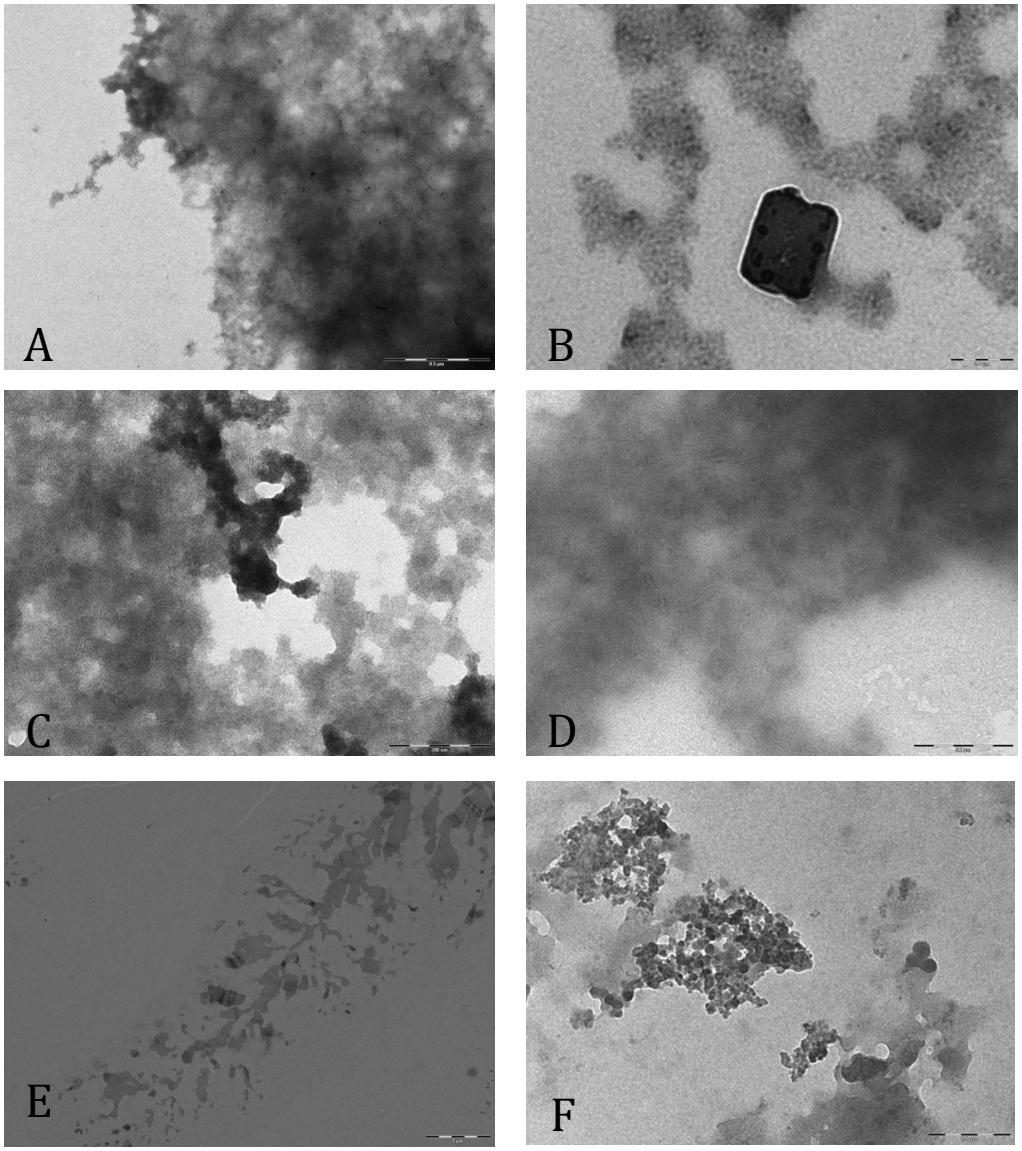


Figure 4.19 TEM images of $(GPO)_7$ peptides in various electrolytes. All images were non-conclusive, except for 150 mM NaCl. Peptide samples were incubated at 8 minutes at 7 mg/ml in their respective buffers. (A) 150 mM NaCl (at 500 nm) at room temperature (B) 150 mM NaCl at 70 °C (scale= 500 nm) (C) Phosphate buffer (at 200 nm) (D) 20mM PBS (at 200 nm) (E) Water (at 1μm) (F) Water at 70 °C (at 200 nm). These results were reproducible.

4.4.5.6 Analysis of $(GPO)_7$ peptides by AFM

150 mM NaCl at pH-7, 20 mM PBS buffer at pH-7, phosphate buffer at pH-7, water, higher salt buffer (213 mM NaCl) at pH-7 and 10 % acetic acid buffer at pH- 2.5 were used for the AFM analysis on $(GPO)_7$ peptides, to observe and analyze the morphologies of supra-assemblies. These specific buffers were chosen because they have favored supra-assembly

formations in collagen and collagen polypeptides in previous studies [48, 60, 109] and also in this research work.

Samples were prepared at 7 mg/ml concentration and incubated for 8 minutes, at room temperature (20 °C) in their respective buffers. Heat-treated samples were prepared at 7 mg/ml, incubated for 8 minutes at 70 °C and cooled down to 20 °C.

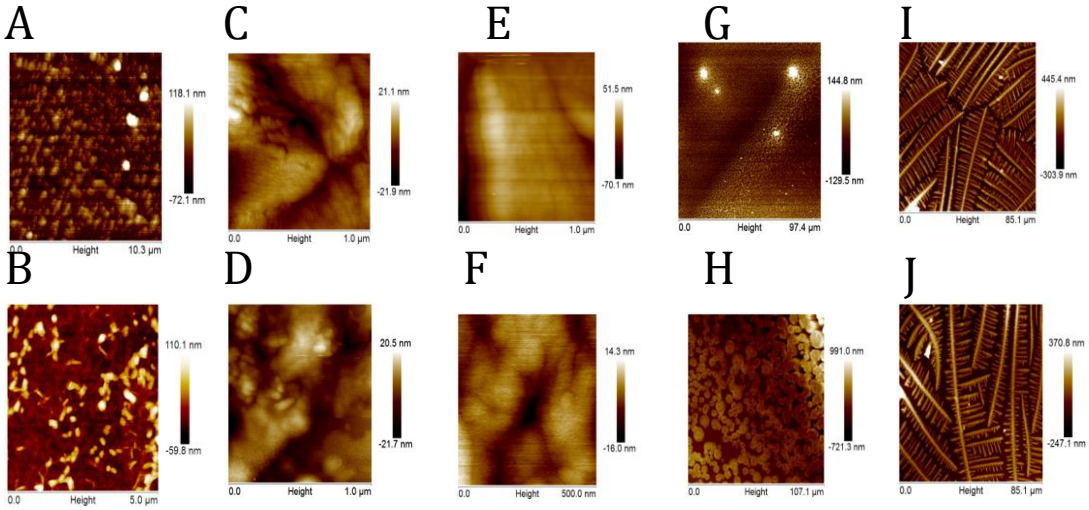


Figure 4.20 Atomic force microscopy imaging of $(GPO)_7$ peptides at 7 mg/ml concentration incubated in (a) water (b) water at 70 °C for 8 minutes (c) 10 % acetic acid (d) 10 % acetic acid at 70 °C for 8 minutes (e) 20 mM PBS buffer (f) 20 mM PBS buffer at 70 °C for 8 minutes (g) phosphate buffer (h) phosphate buffer at 70 °C for 8 minutes (i) Artefacts seen in high salt PBS buffer (213 mM NaCl) (j) Artefacts in high salt PBS buffer at 70 °C for 8 minutes.

AFM analysis in water did not show any fibril formation (Figure 4.20A). Fibrils were formed after the heating process (Figure 4.20B). AFM imaging was carried out on the other unheated samples in 10 % acetic acid (Figure 4.20C), 20 mM PBS buffer (Figure 4.20E), phosphate buffer (Figure 4.20G) and in high salt buffers (Figure 4.20I). AFM analysis was also performed on the heated treated samples shown in Figure 4.20 B, D, F, H and J. For $(GPO)_7$ peptides, decrease in heights after heating was observed for most of the cases.

Detailed cross-sectional analysis and diametrical analysis of self-associated structures have been performed with AFM as shown in Figure 4.21. The reduction in assembly heights for peptides supra-structures is supported by DLS measurement on the aggregations in this chapter.

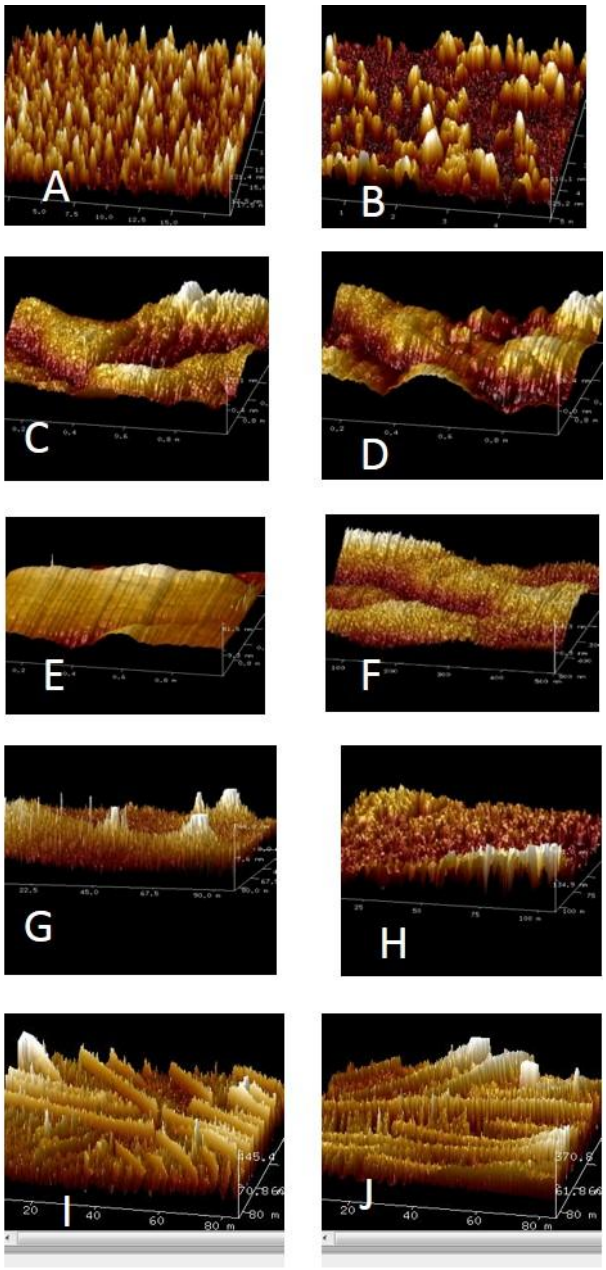


Figure 4.21 3D AFM imaging of $(GPO)_7$ peptides (A) at a scan size of 19.4 µm, unheated $(GPO)_7$ assemblies in water were 121 nm (B) at a scan size of 5 µm, heated treated $(GPO)_7$ assemblies in water were 110 nm, (C) at scan size of 1 µm, unheated acetic acid assemblies were 21.1 nm, (D) at scan size of 500 nm, heated acetic acid assemblies were of 3.1 nm height (E) Scan size of 1 µm in heated 20 mM PBS buffer, heights were 51.5 nm (F) Scan size of 500 nm, unheated 20 mM PBS assemblies = 28.8 nm in (G) Scan size of 97.4 µm in unheated phosphate buffer, assemblies ~ 144 nm (H) Scan size of 107.1 µm, heated phosphate assemblies ~ 991 nm , (I) Scan size of 85.1 µm, unheated high salt PBS assemblies = 445 nm (J) Scan size of 85.1 µm, heated high salt PBS assemblies were 370 nm.

4.5 Summary and findings

From the DLS analysis it was observed that the heat treatment causes reduction in the size of the aggregations. AFM analysis also supports the reduction in heights of the supra-molecular assemblies of collagen mimetic peptides after heat treatments. The shrinkage of these supramolecular assemblies could be a result of the breakage of the intra and inter helical H-bonds after heat treatment.

A decrease in the intermolecular H-bonding for the collagen model peptides is seen by the amide I shift to the higher wavenumbers [148]. The heated supra-molecular assemblies and collagen peptide forming crystals also show peak shifts to higher wavenumbers. Heat treatments results in breaking of H-bond after heating samples. On cooling down, not all the H-bonds would have had re-formed, which would have left behind partially unwound, shrunk helical structures.

$(GPO)_3$ and $(GPO)_5$ peptides were seen to form random aggregations when subjected to heat treatment till 70 °C at a concentration of 7 mg/ml for 8 minutes and on cooling down to 20 °C. $(GPO)_{10}$ and $(GPO)_7$ peptides both favored supramolecular assembly formations under the same heat treatment conditions.

The D-spacing for $(GPO)_{10}$ and $(GPO)_7$ peptides were found to be 12.26 Å, indicating similarities between packing of both triple helices. The D-spacing for the $(GOO)_{10}$ peptides was ~ 12.36 Å that would create a different staggered pattern for the helical arrangement and hence yield a different supramolecular assembly pattern.

5 Crystal structure of $(GPO)_{10}$ and studying racemic $(GPO)_5$-Gly

In this chapter (GPO) collagen model peptides of varying chain lengths have been successfully crystallized and the X-ray structure of $(GPO)_{10}$ has been solved.

5.1 Background

At the atomic scale, the X-ray crystal structure of $(POG)_4(POA)(POG)_5$ peptide provides the first high resolution analysis of a collagen model triple helix at 1.9 Å resolution [169]. This structure shows a 7/2 helical symmetry instead of 10/3 symmetry which was initially proposed for native collagen [170]. A key role of water molecules was also proposed to determine the stability of collagen peptide helix networks.

Berisio et al solved the structure of $(POG)_{10}$ at 1.4 Å resolution [171]. [172] solved the structure of $(POG)_{10}$ at 1.2 Å resolution. Structure of $(GPO)_9$ has been solved at 1.4 Å resolution [173]. $(POG)_{11}$ has been solved at 1.2 Å resolution [172]. In all the aforementioned papers, molecular replacement technique was used to solve the structures.

Although the above structural studies demonstrate common features like triple helical organization of staggered chains stabilized by backbone hydrogen bonds, a few questions still remain unanswered e.g. the conformation of the proline ring, the extent of helix hydration and specific inter-helix interactions, together with the difference in helical pitch with the initial model for native collagen [174].

5.2 Introduction

Successful crystallisation of $(GPO)_{10}$, $(GPO)_7$, NH_2-$(GPO)_7$-NH_2, racemic $(GPO)_5$-Gly and modified peptide like $(GOO)_{10}$ have been reported in this chapter.

The X-ray crystal structure of the $(GPO)_{10}$ peptide, solved at 0.89 Å resolution by direct methods is the most important finding of this chapter. Such sub-Ångstrom resolution of $(GPO)_{10}$ is unprecedented for collagen model peptides.

Table 5.1 List of examples of structures of collagen peptides; solved by X-ray crystallography [81] [174]

	Peptide Sequence	**PDB ID**	**Structural details**
Polytripeptides			
	$(Pro\text{-}Pro\text{-}Gly)_n$ Where n = 9,10	1A3I, 1K6F, 2CUO	7/2 symmetry, NH (Gly)…CO(X)
	$(Pro\text{-}Hyp\text{-}Gly)_{10}$	1V7H	
	$(Gly\text{-}Hyp\text{-}Hyp)_9$	1YM8	
Peptides with collagen sequences:			
LOG2	$(Pro\text{-}Hyp\text{-}Gly)_4$-$(Leu\text{-}Hyp\text{-}Gly)_2$-$(Pro\text{-}Hyp\text{-}Gly)_4$		NH(Gly)…CO(X), NH(X)..W..CO(Gly), 7/2 symmetry in ends, more relaxed in central imino acid poor zone
Integrin binding peptide	$(Gly\text{-}Pro\text{-}Hyp)_2$-Gly-Phe-Hyp-Gly-Glu-Arg-$(Gly\text{-}Pro\text{-}Hyp)_3$	1DZI, 1Q7D	
T3-785 peptide	$(Pro\text{-}Hyp\text{-}Gly)_3$-Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala-Gly-$(Pro\text{-}Hyp\text{-}Gly)_4$	1BKV	
Peptides with Imperfect Gly-X-Y repeat:			
→ Peptide with natural interruption	$(Pro–Hyp–Gly)_3$–Pro–Hyp–Gly–Pro–Gly–$(Pro–Hyp–Gly)_5$	1EI8	Altered H bonding at interruption site, good helix at both ends, but out of register
Peptide with Gly Ala mutation	$(Pro\text{-}Hyp\text{-}Gly)_4$-Pro-Hyp-Ala-$(Pro\text{-}Hyp\text{-}Gly)_5$	1CAG, 1CGD	

This structure offers a new high-resolution basis for discussing the conformation of proline rings, the overall helical symmetry and the role of water molecules in the crystal structure of short peptide sequences, such as GPP/GPO-containing sequences. These sequences mimic the first step of the hierarchical assembly of native collagens. By giving details from a single chain to the pseudo-hexagonal network of triple helices, the structure shows two states of proline rings, a well-defined hydration shell involved in the stabilisation of the inter-helix crystal network, together with helix pitch variation around a model 7/2 helix (3.5 residues per turn).

A 21 residue repeat of GPO peptides has already been described with 7 residue chains each, in its GPO triple helix by Okuyama et al [172] at 1.26 Å. In this chapter, dataset collection at a higher resolution was attempted for $(GPO)_7$ peptide of 21 residues. However, the crystal only diffracted at a resolution of 1.3 Å. A new structure solution was not further obtained with this, because Okuyama et al., have already deciphered the structural details of $(GPO)_7$ at a higher resolution than this.

With the use of conventional methods of crystallisation, $(GPO)_7$ peptides of chain lengths of up to 21 residues have been successfully crystallized. To crystallise shorter GPO residue repeats, racemic crystallisation was performed and successful crystallization of 16 residue $(GPO)_5$-Gly collagen peptide was carried out.

5.3 Result and Discussion

5.3.1 Principle of vapor diffusion and crystal formation

The basic principle of vapor diffusion is the concentration difference in the reservoir and crystallant in the droplet, which drives the water vapor to pass from the droplet to the reservoir. It thus forces the drop to dry and shrink. This increase in protein/peptide concentration forms a crystal [175, 176]. Figure 5.1 (A) explains the phase diagram of vapor diffusion. The factors that affect crystal formation, are (1) rate of reached equilibrium (2) temperature, (3) pH, (4) ionic strength, (5) concentration of precipitant, (6) concentration of biomacromolecules and (7) peptide sequence.

Hanging drop vapor diffusion involves a system where a drop of protein/peptide solution is placed inverted on a cover slip (usually glass) and suspended above the reservoir

as shown in Figure 5.1B [71]. In the sitting drop vapor diffusion method, the drop is placed in a pedestal, which is separated from the reservoir as shown in the Figure 5.1C.

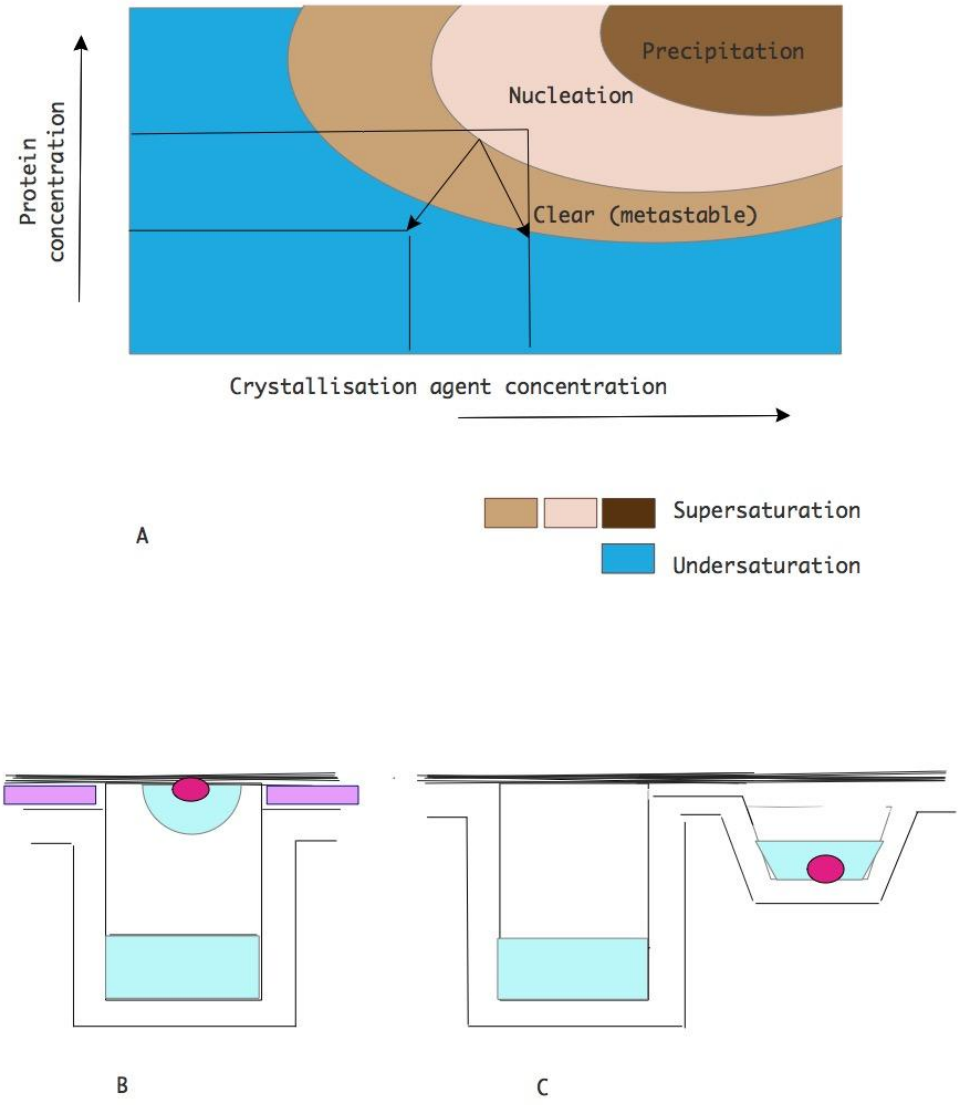


Figure 5.1 (A) Phase diagram for vapour diffusion experiment crystal growing [177]. Vapour diffusion crystallisation methods (B) Hanging drop crystallisation (C) Sitting drop crystallisation. Figures adapted from [178]. Figures drawn in ConceptDraw PRO

5.4 Crystallisation of collagen model peptides

For selected peptides, to form crystals, the GPO/GPP protocols for crystal formations were initially followed [70, 179]. These were later modified. For example, in the literature the $(GPO)_{10}$ peptide has been previously crystallised at a concentration of 4.5 mg/ml at 10 % acetic acid solution and 22 % PEG 200 [172]. In our research, $(GPO)_{10}$ and $(GPO)_7$ peptides formed crystals while the shorter chain length of peptides, $(GPO)_5$ and $(GPO)_3$ formed only precipitates (see section 5.4.2). Apart from the NH_2-$(GPO)_n$-COOH peptides, NH_2-$GPO)_7$-NH_2 and $(GPO)_5$-Gly also formed crystals.

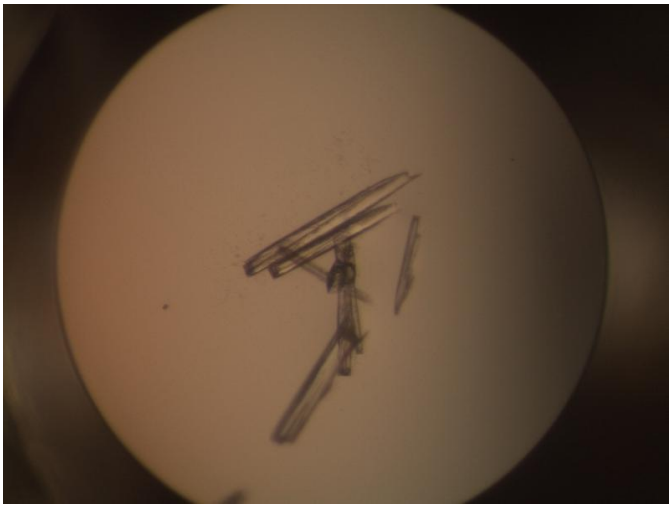
5.4.1 Peptide concentration proved inversely proportional to (GPO) chain lengths for crystal formation

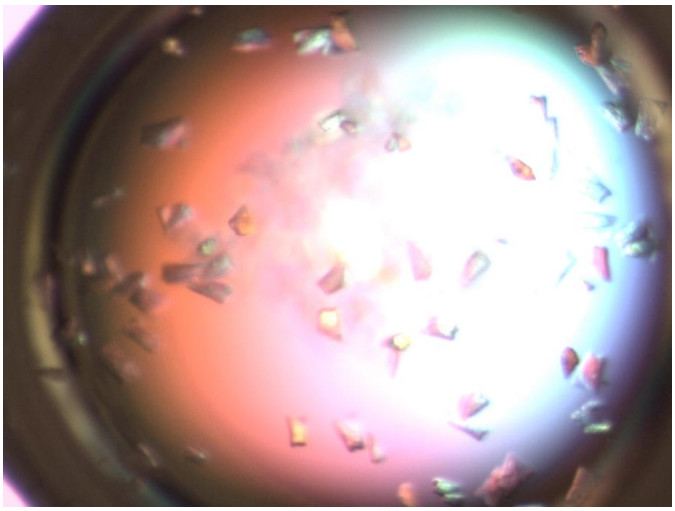
Peptide concentration screens were prepared from 1 mg/ml to 20 mg/ml concentrations in 10 % acetic acid solutions. Shorter chain length peptides like $(GPO)_3$ and $(GPO)_5$ did not form crystals but showed prolonged precipitations at very high concentrations of 15 mg/ml and 20 mg/ml. Crystallization was achieved for the next set of longer chain lengths of peptides, NH_2-$(GPO)_7$-NH_2 and $(GPO)_7$ at 15 mg/ml.

For the longest chain lengths of peptides of 30 residue chain lengths, crystals were made at low concentrations. $(GOO)_{10}$ formed crystals at 7 mg/ml concentration, while $(GPO)_{10}$ formed crystals at only 5 mg/ml concentration. At consistent conditions, the sequence make up of Gly-Pro-Hyp was found to form crystals at lower concentrations, whereas the sequence make up of Gly-Hyp-Hyp repeats formed crystals at slightly higher peptide concentrations (7 mg/ml) for the same peptide chain lengths. $(GPO)_5$-Gly peptides was a special case that formed crystals at both lowest and highest concentrations of 7 mg/ml and 15 mg/ml respectively.

5.4.2 GPO crystal formations

All crystals were obtained in 10 % acetic acid buffer at pH 2.4. $(GOO)_{10}$ peptides formed crystals at 7 mg/ml concentration in 10 % acetic acid solution in 22 % PEG 400 at incubation temperatures of 20 °C and also at 4 °C (see Figure 5.2 A). Heated crystals used as samples for IR spectroscopy from chapter 4 were also obtained at the same conditions.

A 

B 

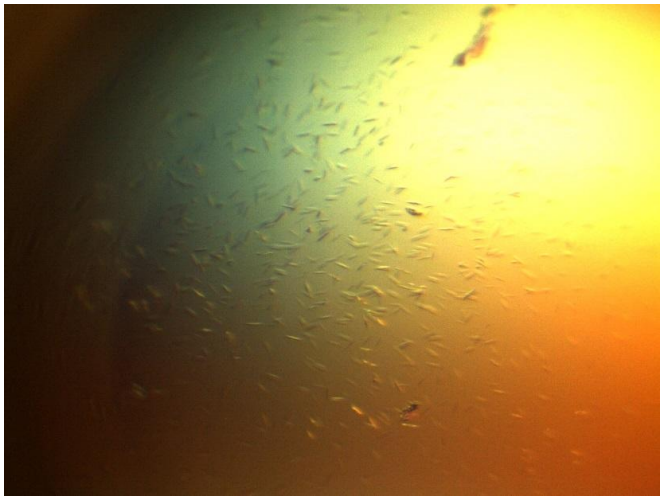
C 

Figure 5.2 Crystallisation of GPO peptides (A) $(GOO)_{10}$ as needle shaped crystals (B) $(GPO)_7$ crystallised as square and rectangle shaped crystals (C) NH_2-$GPO)_7$-NH_2 as rice shaped crystals

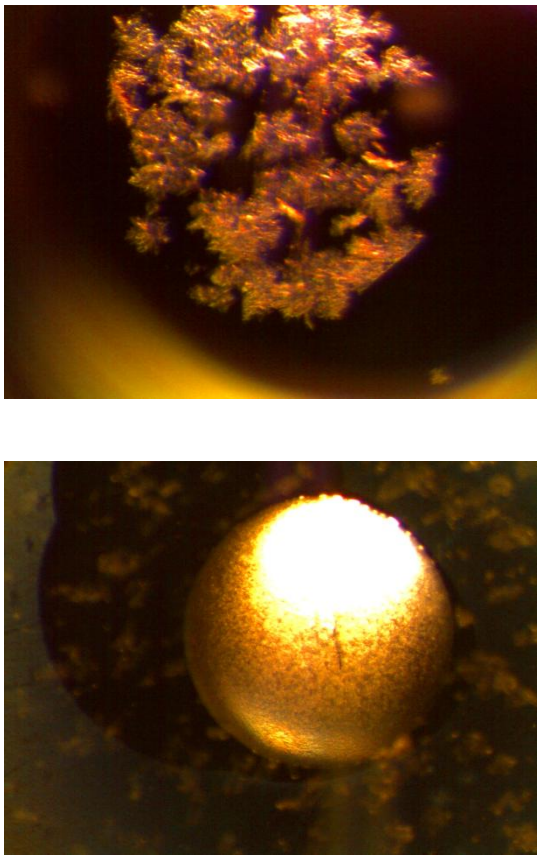


Figure 5.3 Precipitate formation (A) $(GPO)_5$ crystals (B) $(GPO)_3$ crystals

$(GPO)_7$ formed crystals at 15 mg/ml concentration at 30 % and 40 % PEG 400 (see Figure 5.2B). NH_2-$(GPO)_7$-NH_2 formed crystals (Figure 5.2C) at PEG conditions of both 30 % PEG 550 and at 58 % PEG 300. These were achieved at 4 °C, at concentrations of 10 mg/ml and also 15 mg/ml. $(GPO)_{10}$ crystals formed at 4 mg/ml, 5 mg/ml, 7 mg/ml and 15 mg/ml concentrations at 4 °C with 22 % PEG 400 (see Figure 5.11).

$(GPO)_5$ formed crystalline precipitates at concentrations of 15 mg/ml and 20 mg/ml with 35 % PEG 400 at an incubation time of 1 week at 4 °C (see Figure 5.3A). $(GPO)_3$ formed precipitates mostly at 15 mg/ml and some precipitation also occurred at 20 mg/ml concentration with PEG 400 at percentages of 22 %, 25 %, 30 % and 35 % (see Figure 5.3B). These formed at 4 °C and also at a room temperature of 20 °C.

Racemic $(GPO)_5$-Gly peptides formed crystals at 15 mg/ml concentration and also at 7 mg/ml concentration at 58 % PEG 300 at both 4 °C and also at 20 °C. Their incubation times were 2 to 3 weeks (see Figure 5.4). Along with the above collagen model peptide

crystals, their heated forms were prepared in the same conditions as per infrared spectroscopic analysis in chapter 4.

5.5 Racemic $(GPO)_5$-Gly

Racemic peptides were selected for the crystallisation of the smaller length of peptides. $(GPO)_5$-Gly peptides were selected instead of $(GPO)_5$ because of the ease of its synthesis, by the Maurice Wilkins Centre.

Successful racemic crystallisation of the smallest possible collagen model peptide of 16 residues, $(GPO)_5$-Gly has been achieved that diffracted at both 1 Å and 1.5 Å resolutions. However, the structure could not be solved (The data processing section from chapter of materials and methods explains the detailed procedures used to solve the structure).

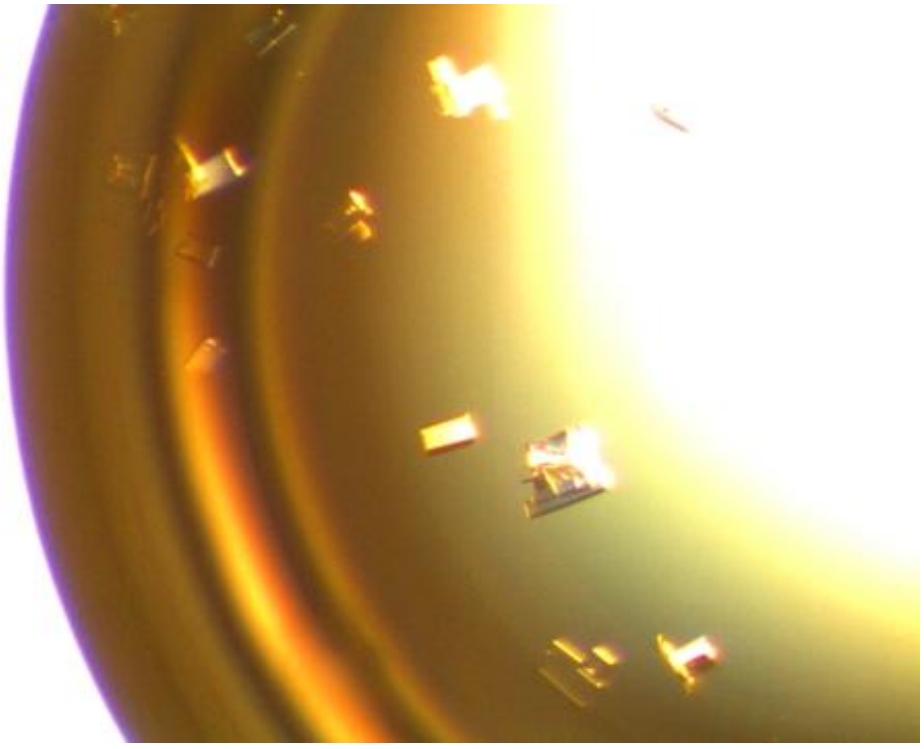


Figure 5.4 Racemic crystallisation of $(GPO)_5$-Gly peptides; Crystals formed at 15 mg/ml concentrations and also at 7 mg/ml concentrations at 58 % PEG 300 at 4 °C

5.5.1 CD analysis of D and L $(GPO)_5$-Gly peptides;

Both L and D forms of $(GPO)_5$-Gly peptides were studied for the secondary structure formation [48].

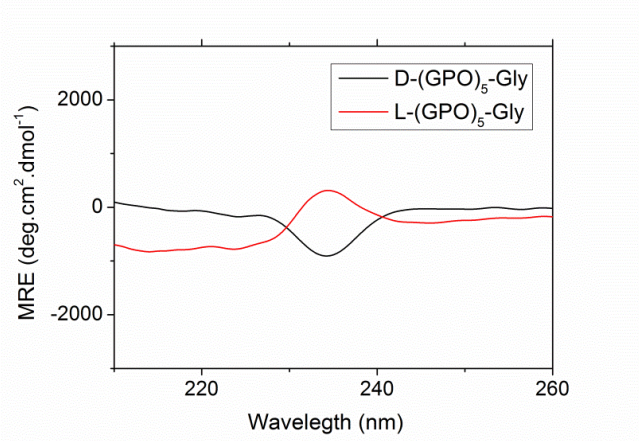


Figure 5.5 Analysis by CD; Secondary structure determination with D and L $(GPO)_5$-Gly peptides at 1 mg/ml in 20 mM PBS buffer, pH-7.

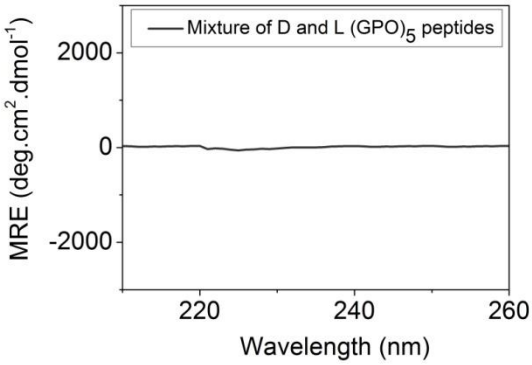


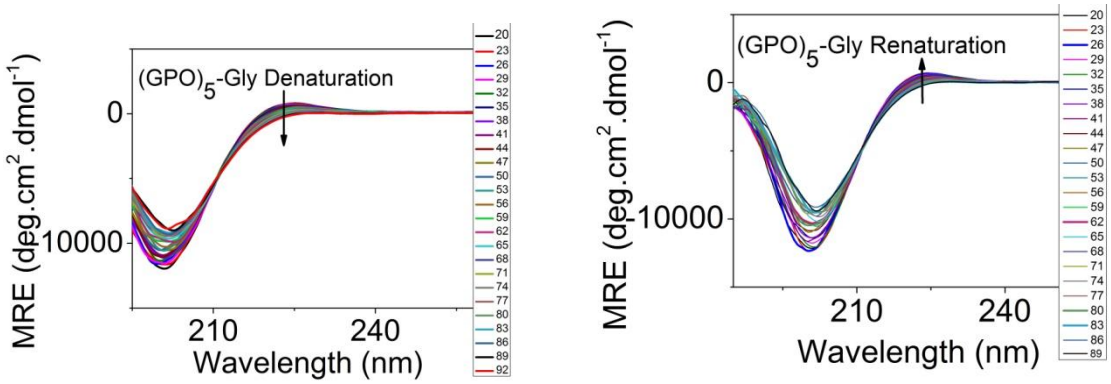
Figure 5.6 Equal proportions of D and L $(GPO)_5$ peptides shows a blank read

With the L form, triple helices were seen (maximum at 225 nm wavelength whereas with the D form, an absolute opposite minimum was seen at 225 nm. The structure was uncertain (see Figure 5.5). On mixing both peptides i.e. a racemic mixture of 1:1 ratio of D and L peptides, a blank reading with no signals was obtained. This is the usual expectation, when racemic mixtures are analysed.

5.5.2 Manual melting analysis of $(GPO)_5$-Gly triple helices by CD

L-$(GPO)_5$-Gly peptide samples were prepared at 1 mg/ml in 20 mM PBS buffer, pH-7 and were incubated at 4 °C for 2 days. There was no wavelength shift at 200 nm and the minimum (presence of random coils) stayed intact throughout temperature increments (Figure 5.7A) and decrements (Figure 5.7B).

The melting point for $(GPO)_5$-Gly was found to be ~ 35 °C.



A B

Figure 5.7 L-$(GPO)_5$-Gly peptides at 1 mg/ml in 20 mM PBS buffer, pH-7; Analysis by CD (A) Manual unwinding L-$(GPO)_5$-Gly peptides (B) Manual re-winding

5.5.3 Melting analysis of L-$(GPO)_5$-Gly triple helices by DSC

DSC analysis on $(GPO)_5$-Gly peptides in Figure 5.8 shows no triple helical peak at 5 mg/ml concentration of peptides. No further lesser concentrations of peptides were used for examination. The reason behind absence of triple helical peak at a high concentration of 5 mg/ml is believed to be the short chain length of $(GPO)_5$-Gly. In chapter 3, it was seen that peptides with chain lengths shorter than $(GPO)_7$ do not show DSC melting peaks even at high concentrations like 5 mg/ml.

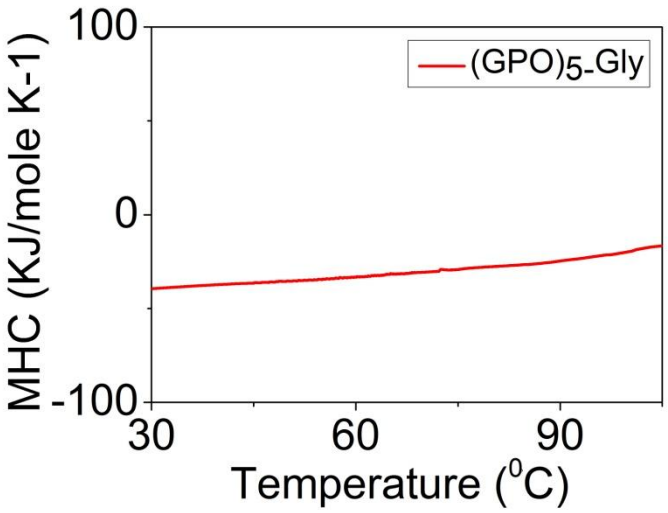


Figure 5.8 DSC Analysis; Studying melting for triple helices formed by L-$(GPO)_5$-Gly peptides at 5 mg/ml in 20 mM PBS buffer

5.5.4 Supramolecular assembly formation;

5.5.4.1 Impact of heat on L-$(GPO)_5$-Gly peptides; Analysis by DLS

Supramolecular species of diameter 140 nm and 414 nm were seen for the unheated $(GPO)_5$-Gly peptides(Figure 5.9A). Supramolecular assemblies of decreased sizes of ~ 138 nm were observed for the heat-treated samples (heated at 70 °C incubation for 8 minutes and cooled down to 20 °C) as shown in Figure 5.9B. Heating brings about reduction in the size of supramolecular assemblies (as seen in chapter 4)

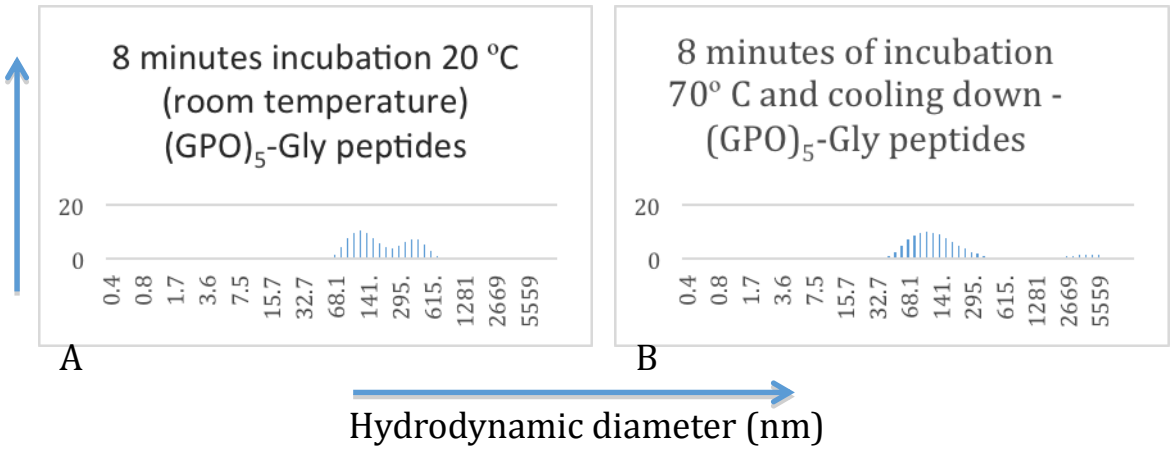


Figure 5.9 DLS analysis (intensity vs hydrodynamic diameter) for L-$(GPO)_5$-Gly peptides at 7 mg/ml in 20 mM PBS buffer, pH 7 (A) unheated peptides incubated at room temperature for 8 minutes (B) heat treated peptides at 70 °C incubated for 8 minutes and cooled down to 20 °C

5.5.4.2 Impact of heat on L-$(GPO)_5$-Gly peptides; Analysis by TEM

In TEM analysis of L-$(GPO)_5$-Gly peptides (Figure 5.10A), for the unheated peptides incubated at room temperature for 8 minutes, no clear supramolecular assemblies were observed. For the heat-treated peptides (70 °C incubation for 8 minutes and cooling down to 20 °C (Figure 5.10B)) aggregations were noticeable. No significant differences in the size of assemblies could be confirmed by TEM for the unheated and the heat- treated peptides.

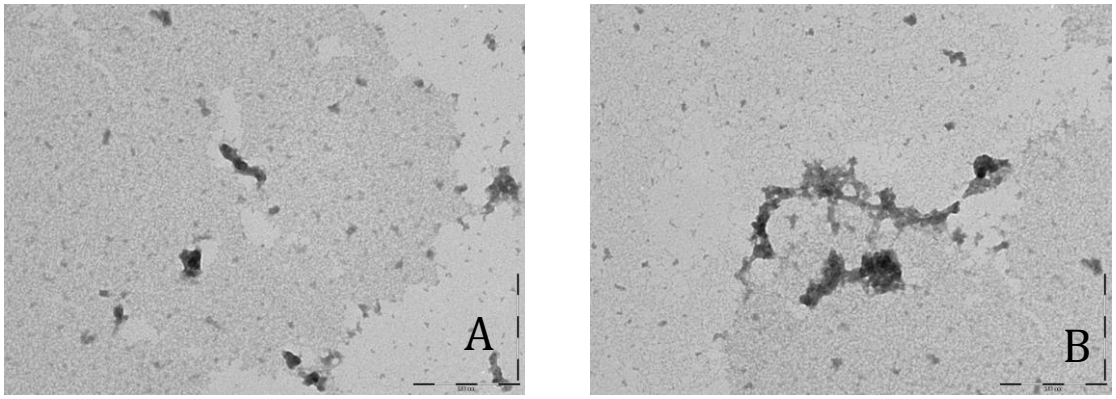


Figure 5.10 Analysis by TEM; Supramolecular assembly formations for L-$(GPO)_5$-Gly peptides at 7 mg/ml in 20 mM PBS buffer, pH 7 (A) for unheated peptides incubated at room temperature for 8 minutes, no clear supramolecular assemblies were noticed (B) after heat treatment at 70 °C incubation for 8 minutes and on cooling down, aggregations were noticed. TEM artefacts were differentiated from assemblies based on following references [131], [132].

5.6 Structure of $(GPO)_{10}$

5.6.1 Structure determination and refinement

The $(GPO)_{10}$ crystals have been obtained at a concentration of 4 mg/ml at 4 °C within two weeks of incubation time for X-ray diffraction.The diffraction pattern has been obtained at a resolution of 0.89 Å at Australian synchrotron in collaboration with Dr Hironori Suzuki.

5.6.2 Structure solution

Structure was solved by Dr Hironori Suzuki. In the MX2 beamline, 1440 images were collected for $(GPO)_{10}$ peptides. With $XD_{S,}$ the images were integrated in the spacegroup P1. From among the possible space groups and cell dimensions from the CORRECT.LP file, the spacegroup and cell dimensions as shown in Table 5.2 were selected. These parameters were updated in the generated XDS.INP file. Scaling of the data was conducted in XDS. The scaled .HKL file was converted into an SHELX.HKL file in XDS. This was the conversion step for preparation of an input file for running SHLEXD. Along with SHELX.HKL file, the other input files for SHELXD were the .ins file with the same cell parameters and the .res file. The model generated from SHELXD was further built. Refinements were performed with the scaled .HKL file in REFMAC5 alongside building the structure. This was based on the visibility of proline and hydroxyproline rings. After 26 refinements, a sharp decrease in R_{work} / R_{free} was not$_{iced}$ aft$_{er\ t}$he introduction of water molecules, towards the end stages of refinements. The final structure determining R_{work} / R_{free} (%) was 6.8 / 7.7 (see Table 5.2).

After 26 refinements, this solved structure had alpha carbons of prolines and hydroxyprolines in their R configurations. Thus the structure was inverted to get the opposite enantiomer in the final step, because the natural configuration of alpha carbons in prolines and hydroxyprolines is S. The final inverted structure has been discussed below.

5.6.3 General structure description

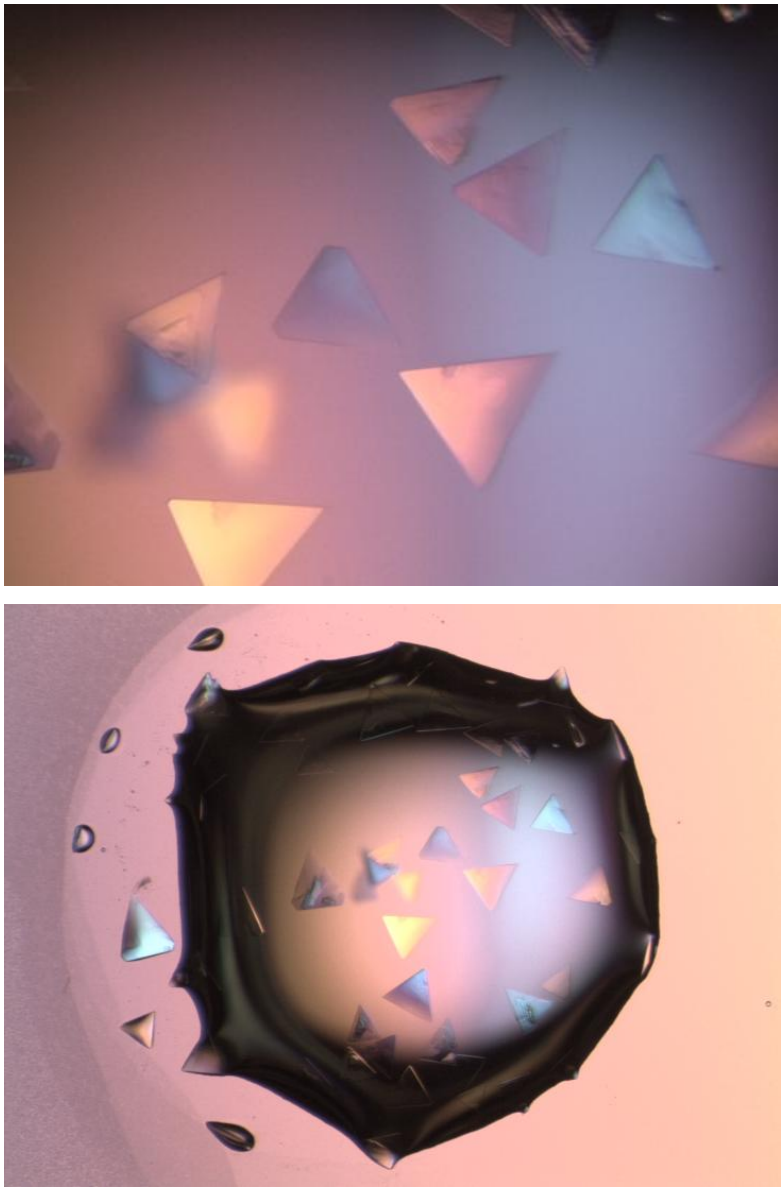


Figure 5.11 Crystals of $(GPO)_{10}$ peptide crystallised as triangular crystals in the presence of acetic acid and polyethylene glycol (at both 4 mg/ml and 5 mg/ml concentration, PEG 400)

The asymmetric unit comprises of one chain of 21 amino acid residues (7 GPO triads) and 52 water molecules. Structural details have been shown in Figure 5.12. Glycine residues have been shown in grey, proline in green and hydroxyproline in red. In this structure of triple helix forming $(GPO)_{10}$ peptide, the triple helical packing has also been in P21.

From previous studies on the triple helix forming peptides like $(PPG)_4$-POG-$(PPG)_4$, $(PPG)_4$-OOG-$(PPG)_4$, $(GPO)_7$ etc, the packing has been also been seen in the P21 symmetry [114, 180].

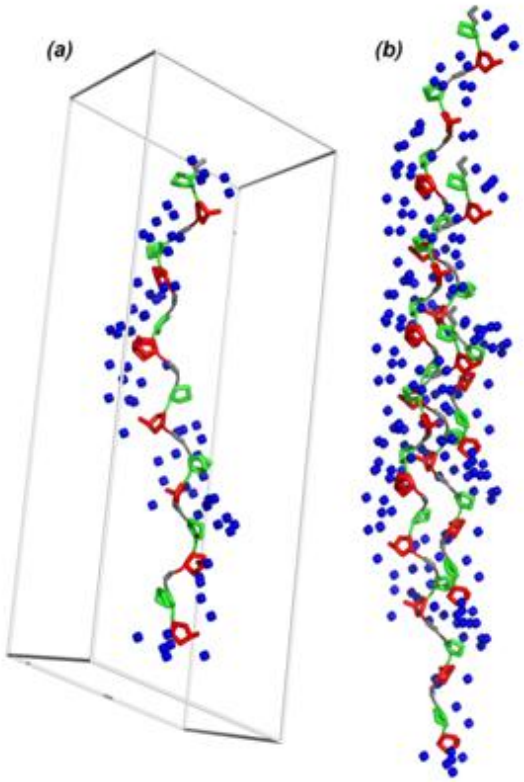


Figure 5.12 (1) Crystal structure of $(GPO)_{10}$ at 0.89 Å (glycine: grey, proline: green, hydroxyproline: red, water molecules: blue). (A) Asymmetric unit comprising a single chain of 21 residues and 52 water molecules; (B) triple helix side view; This has been generated in Pymol.

The water molecules have been shown in blue in Figure 5.12A. The triple helix side view is shown in Figure 5.12B. The pseudo-hexagonal network of triple helices is shown in Figure 5.13. Table 5.2 shows the collection and refinement statistics.

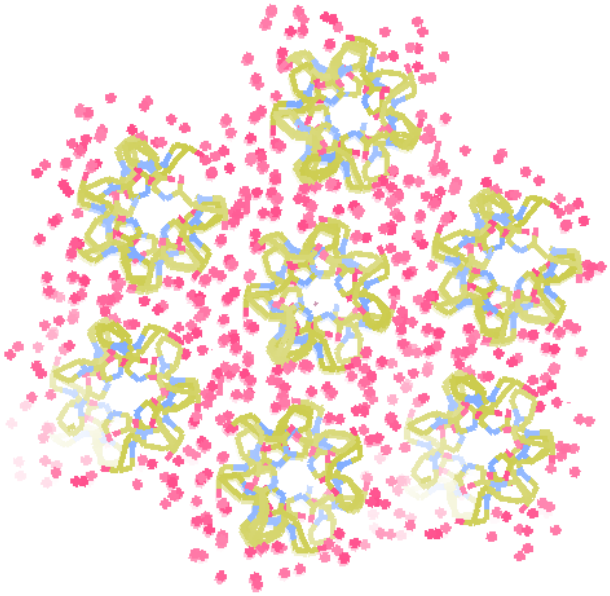


Figure 5.13 Pseudo-hexagonal network of triple helices at 0.89 Å surrounded by pink water molecules. Image was generated in Coot.

$(GPO)_{10}$ crystals have also been re-examined with mass spectrometry for the confirmation of the presence of $(GPO)_{10}$ peptide sequence. From structural analysis, despite the high resolution of the structure, nine residues per peptide chain have not been resolved together with both peptide termini, similar to the previously reported crystal structures of $(POG)_{10}$ and other collagen model peptides [169, 171, 172].

The packing of the molecules in the asymmetric unit shows that there is not enough space to account for the other nine residues of the 30 amino acid long peptide, suggesting that they are not disordered. Instead, this chapter proposes that the unresolved residues are averaging across the peptide chain in the asymmetric unit, given the repeating GPO architecture. This results in a model that behaves as an infinite helix. Since the electron density of the peptide terminals differs only slightly from the one of the backbone amide bonds, the resulting averaged structure apparently lacks both terminals. As commonly obtained for collagen model peptides, three peptide chains are found to be supercoiled around a common axis (Figure 5.12B) with single chains staggered by five residues along the helix axis in our case (Figure 5.12B). These generate triple helices (Figure 5.12C) arranged in a highly hydrated pseudo-hexagonal network, Figure 5.12D).

Table 5.2 Data collection and refinement statistics.

Data collection statistics	
Wavelength (Å)	0.7293
Space group	$P2_1$ (P 1 2_1 1)
Cell dimensions	
a, b, c (Å)	13.9, 26.2, 19.9
α, β, γ (°)	90, 105.4, 90
Resolution (Å) (outer shell)	19.19-0.89 (0.91-0.89)
Number of measured reflections	74,888 (11,554)
Number of unique reflections	10,270 (1,610)
Completeness (%)	96.5 (93.9)
R_{merge} (%)	2.9 (7.7)
$I/\sigma I$	54.7 (25.4)
Refinement statistics	
R_{work} / R_{free} (%)	6.8 / 7.7
Mean B value (Å^2)	2.97
RMSD bond length (Å)	0.047
RMSD bond angle (°)	4.406
Number of molecules in the asymmetric unit	1

Integration-XDS

Solution- ShelxD

Refinement- Refmac5

5.6.4 Conformations of proline and hydroxyproline rings

The pyrrolidine ring of the proline residue can adopt two distinct "puckered" conformations [181]. The ring conformations differ by the Cγ atom and the CO group of the proline residue that lie on the same or opposite side of the plane defined by Cδ, N, and Cα [181].

The conformation of the proline rings in collagen model peptide has been a source of debate in literature [69, 172, 182]. The weak electron density obtained around proline residues generally hampers the complete resolution of the ring conformation. In our structure, the electron density is sufficiently resolved to display two stable states for the rings of proline 5, 14 and 20 (Figure 5.14). The two conformations correspond to two different ring "puckering" states.

This crystal structure shows that the proline rings can adopt two stable conformations in a single crystal of the collagen model peptide, while all hydroxyproline rings remain in the same "up-puckered" conformation. These results are consistent with a reported analysis of the Cambridge Structural Database showing that proline rings can adopt two states, while hydroxyproline rings were found in an "up-puckered" state only. This conformational preference of hydroxyproline rings has been proposed to correlate with the increased helix stability for hydroxyproline-containing collagen model peptides over peptides containing proline only, via favoring the peptide backbone helical fold [61].

5.6.5 Triple helix symmetry

Variation in triple helix pitch has been proposed to play a role in the functional binding of collagen to other biomolecules, such as integrin or heparin [183-185]. The unit cell dimension along the helix axis (19.9 Å) is consistent with an average 7/2 helical pitch (20.0 Å axial repeat, 3.5 residues per turn) [186]. This value is similar to the smallest unit cell that was reported for $(POG)_{10}$, which is a = 13.89 Å, b = 26.12 Å, c = 19.95Å and α = 90, β = 105.95 and γ = 90. [172].

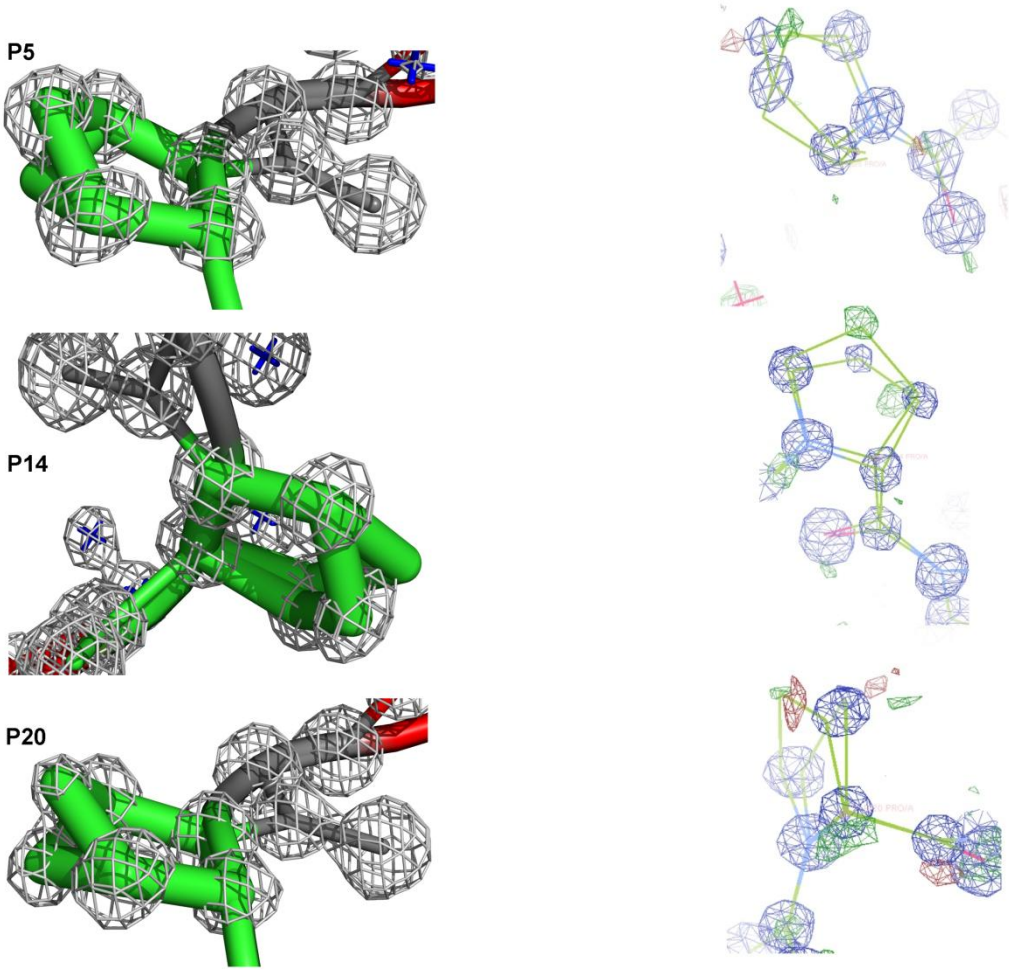


Figure 5.14 Two different conformations for proline 5, 14 and 21. Figure on the left shows the grey mesh corresponds to $2F_0$-F_c map contoured at 1.0 sigma. The figures on the right show the two proline conformations at 3.1 sigma, where green density is the positive density. These proline puckering maps have been generated after the final stages of structure refinement. The structure was built with the model generated from the direct method, SHLEXD. These maps have been generated after the final stage of refinement.

Table 5.3 Average torsion angles for 7 GPO triads in the asymmetric unit compared to the ones obtained for $(POG)_{10}$ crystal structures. These are all 7/2 helices. Standard of deviations are shown in brackets.

G	**φ**	**ψ**	**ω**
$(GPO)_{10}$	**-70.8 (3.4)**	**174.8 (5.9)**	**-179.4 (6.0)**
$(POG)_{10}$[a]	-74.8 (5.6)	172.8 (3.0)	179.2 (1.7)
$(POG)_{10}$[b]	-70.4 (3.8)	174.9 (3.4)	/
$(POG)_{10}$[c]	-71.3 (1.6)	174.2 (1.2)	178.8 (1.1)
P	**φ**	**ψ**	**ω**
$(GPO)_{10}$	**-72.0 (14.6)**	**162.3 (14.4)**	**-171.5 (7.0)**
$(POG)_{10}$[a]	-72.7 (3.5)	161.6 (5.1)	179.6 (2.1)
$(POG)_{10}$[b]	-69.8 (7.9)	162.0 (3.9)	/
$(POG)_{10}$[c]	-71.3 (1.4)	161.5 (1.1)	172.3 (1.0)
O	**φ**	**ψ**	**ω**
$(GPO)_{10}$	**-57.3 (5.0)**	**150.2 (6.3)**	**-175.4 (7.7)**
$(POG)_{10}$[a]	-58.4 (4.8)	152.0 (6.4)	178.5 (2.1)
$(POG)_{10}$[b]	-57.4 (2.0)	149.8 (2.9)	/
$(POG)_{10}$[c]	-56.9 (1.3)	150.0 (1.1)	174.7 (1.1)

[a] from Nagarajan et al. 1999

[b] from Berisio et al. 2001

[c] from Okuyama et al. 2005

Table 5.4 Main chain torsion angles in the asymmetric unit of 7 GPO triads. Phi (φ, deg) is indicative of C'-C', psi (ψ, deg) of N-N, and omega (ω, deg) of Cα-Cα. Proline for which 2 states of ring puckering have been obtained has been marked with *. To gain insights into potential pitch variation, the torsion angles were determined for each residue of the 7 GPO triads in the asymmetric unit. These were compared to the torsion angle values for ideal 7/2 and 10/3 helices (Table 3). Values for ideal 7/2 and 10/3 helices have been established in the last rows of the table. These are from Okuyama et al. 2006. Highest and lowest values are underlined for each type of residue, glycine (G), proline (P) and hydroxyproline (O).

G	φ	ψ	ω	**P**	φ	ψ	ω	**O**	φ	ψ	ω
G1	/	168.7	-177.2	P2	-74.7	171.4	-174.1	O3	-60.5	154.0	177.2
G4	-73.1	180.0	179.5	P5*	-68.3	163.2	-168.0	O6	-55.6	148.7	-170.1
G7	-68.1	170.6	-173.4	P8	-57.7	148.0	-166.4	O9	-52.0	149.6	178.0
G10	-71.3	174.2	-178.5	P11	-76.4	165.7	-178.5	O12	-58.3	152.8	-167.7
G13	-67.2	176.5	175.6	P14*	-78.8	159.7	-169.0	O15	-52.2	144.0	179.0
G16	-72.0	177.1	178.6	P17	-73.2	166.2	-174.1	O18	-62.9	149.4	-173.7
G19	-72.1	177.2	180.0	P20*	-75.7	161.4	-170.1	O21	-55.9	/	/
G-7/2	***-75.7***	***176.3***	***173.5***	***X-7/2***	***-77.9***	***166.1***	***175.8***	***Y-7/2***	***-60.2***	***163.4***	***179.7***
G-10/3	***-78.5***	***147.1***	***-165.6***	***X-10/3***	***-67.6***	***147.7***	***-171.5***	***Y-10/3***	***-69.0***	***155.4***	***-166.9***

The average torsion angles phi (φ, indicative of C'-C'), psi (ψ, indicative of N-N), and omega (ω, indicative of Cα-Cα) are closer to a 7/2 pitch than to the 10/3 native collagen one, as previously observed for other similar crystal structures (Table 5.3 and Table 5.4).

To gain insights into potential pitch variation, the torsion angles were determined for each residue of the 7 GPO triads in the asymmetric unit and compared to the values for ideal 7/2 and 10/3 helices (Table 5.4). The values for 7/2 and 10/3 helices have been shown in bold in Table 5.4. The largest angle variation is observed around proline residues, with differences up to 21.1° for phi (φ), 23.4° for psi(ψ), and 12.1° for omega(ω), from one proline residue to another. These results support the existence of a range of stable helical pitches for the collagen model peptide main chain, around the average 7/2 symmetry previously reported for similar crystal structures. The present structure is consistent with the hypothesis of collagen pitch variation, with a preferred 7/2 pitch for proline-rich domains [33, 179, 187, 188].

5.6.6 Intra- and inter-strand backbone hydrogen bonds

The high resolution of the present crystal structure allows the precise localisation of the hydrogen bond networks involved in the stabilisation of the triple helix and the comparison of these with current models. Within one triple helix, intra-strand and inter-strand backbone hydrogen bonds are detected from the distances between backbone amine groups (NH), carbonyl groups (C=O), and hydroxyl groups from hydroxyproline residues (Table 5.5).

Considering the 21 residues single chain of the asymmetric unit, 7 intra-strand hydrogen bonds could be located. The intra-strand network that contributes to the stabilization of a single helical chain is made of a ladder of hydrogen bonds that involves backbone carbonyl (C=O) and amine (NH) groups around all three types of residue, with no specific discrimination for a specific side-chain environment (Table 5.5, chain A/chain A).

In the inter-strand network stabilising one triple helix, backbone carbonyls (C=O) from both proline and glycine residues are H-bond acceptors, while only backbone amine moieties from glycine residues serve as donors (NH) (Table 5.5). The absence of side-chain steric hindrance around glycine appears required in this H-bond network, while glycine residues form the core of the triple helix.

Table 5.5 Intra- and inter-strand hydrogen bond networks (2Å<distances<4Å) between backbone carbonyls (CO) and backbone amines (NH). This demonstrates two things (1) The intra-strand network that contributes to the stabilization of a single helical chain is made of a ladder of hydrogen bonds that involves backbone carbonyl (C=O) and amine (NH) groups around all three types of residue. (2) Hydroxyproline residue backbone amide groups are not involved in triple helix stabilisation. Along with a side-chain configuration, hydroxyproline rings are correctly exposed, for being involved in interactions between triple helices. This stabilises the pseudo-hexagonal network formed by the triple helices.

	Chain A	***Chain B***	***Chain C***
Chain A	P2(CO)-G4(NH): 3.62Å P5(CO)-G7(NH): 3.52Å P8(CO)-G10(NH): 3.56Å O9(CO)-G10(NH): 2.31Å P11(CO)-G13(NH): 3.60Å P14(CO)-G16(NH): 3.39Å P17(CO)-G19(NH): 3.56Å	G7(CO)-G1(NH): 3.59Å P8(CO)-G1(NH): 3.82Å G10(NH)-P2(CO): 2.85Å G10(CO)-G4(NH): 3.73Å P11(CO)-G4(NH): 3.79Å G13(NH)-P5(CO): 2.87Å G13(CO)-G7(NH): 3.78Å P14(CO)-G7(NH): 3.97Å G16(NH)-P8(CO): 2.83Å G19(NH)-P11(CO): 2.97Å G19(CO)-G13(NH): 3.74Å	P2(CO)-G10(NH): 2.85Å P5(CO)-G13(NH): 2.84Å G7(NH)-G13(CO): 3.78Å P8(CO)-G16(NH): 2.83Å P11(CO)-G19(NH): 2.97Å G13(NH)-G19(CO): 3.74Å G13(NH)-G20(CO): 3.96Å
Chain B	/	/	G1(NH)-P14(CO): 3.03Å G4(NH)-P17(CO): 2.97Å G4(CO)-G19(NH): 3.89Å G7(NH)-P20(CO): 2.87Å
Chain C	/	/	/

This glycine-stabilised inter-strand network is consistent with the current model structure for native collagen, which also displays a glycine residue as the core of the triple helix [186]. The fact that the hydroxyproline residue backbone amide groups are not involved in triple helix stabilisation (intra-strand interactions only, Table 5.5, Figure 5.15) is also consistent with this collagen model displaying hydroxyproline side-chains exposed on the outside of the triple helix. With this side-chain configuration, hydroxyproline rings are correctly exposed for being involved in interactions between triple helices, to stabilise the pseudo-hexagonal network they form.

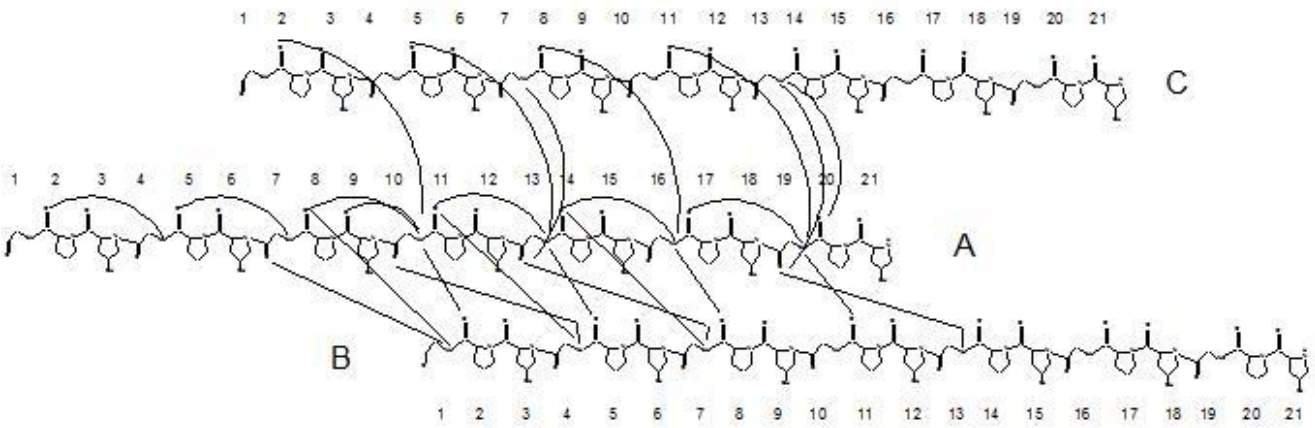


Figure 5.15 Diagrammatic representation of intra- and inter-strand hydrogen bond networks (2Å<distances<4Å) from Table 5.5, between backbone carbonyls (CO) and backbone amines (NH)

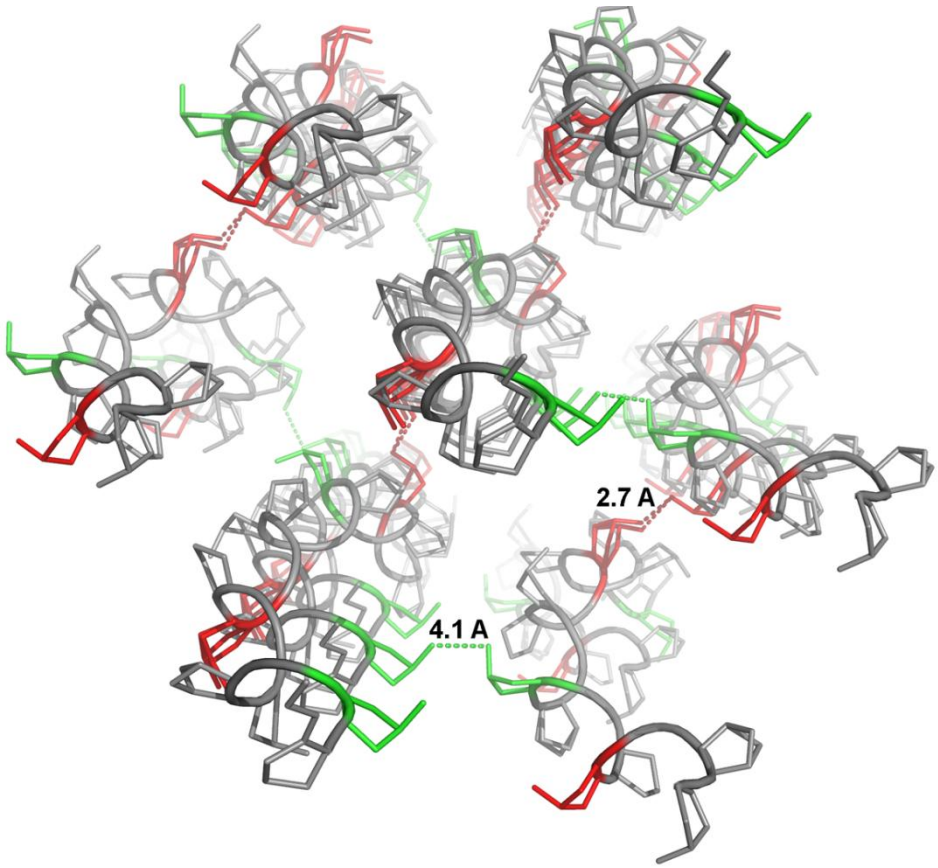


Figure 5.16 Pseudo-hexagonal network of triple helices showing close distances for two pairs of hydroxyproline: 4.1 Å for Hyp3-Hyp15 and 2.7 Å for Hyp9-Hyp18.

5.6.7 Interactions between helices and helix hydration

The set of interactions stabilizing the crystalline network of triple helices has been the subject of debate, especially regarding the role of hydroxyproline residues and water molecules. The

$(GPO)_{10}$ triple helices here define a highly hydrated pseudo-hexagonal network (see Figure 5.13).

Short atomic distances (2.7-4.1 Å) are found between one helix and four neighboring triple helices, between the hydroxyl groups of two specific couples of hydroxyproline residues, respectively Hyp9-Hyp18 (2.7 Å) and Hyp3-Hyp15 (4.1 Å) (Figure 5.16). No density was seen for the hydrogens. Although the first distance value, 2.7 Å is consistent with direct hydrogen bonding, the second, 4.0 Å is too long to support such a direct interaction (hydrogen bonding distance between O and O should be 2.8-3.2Å).

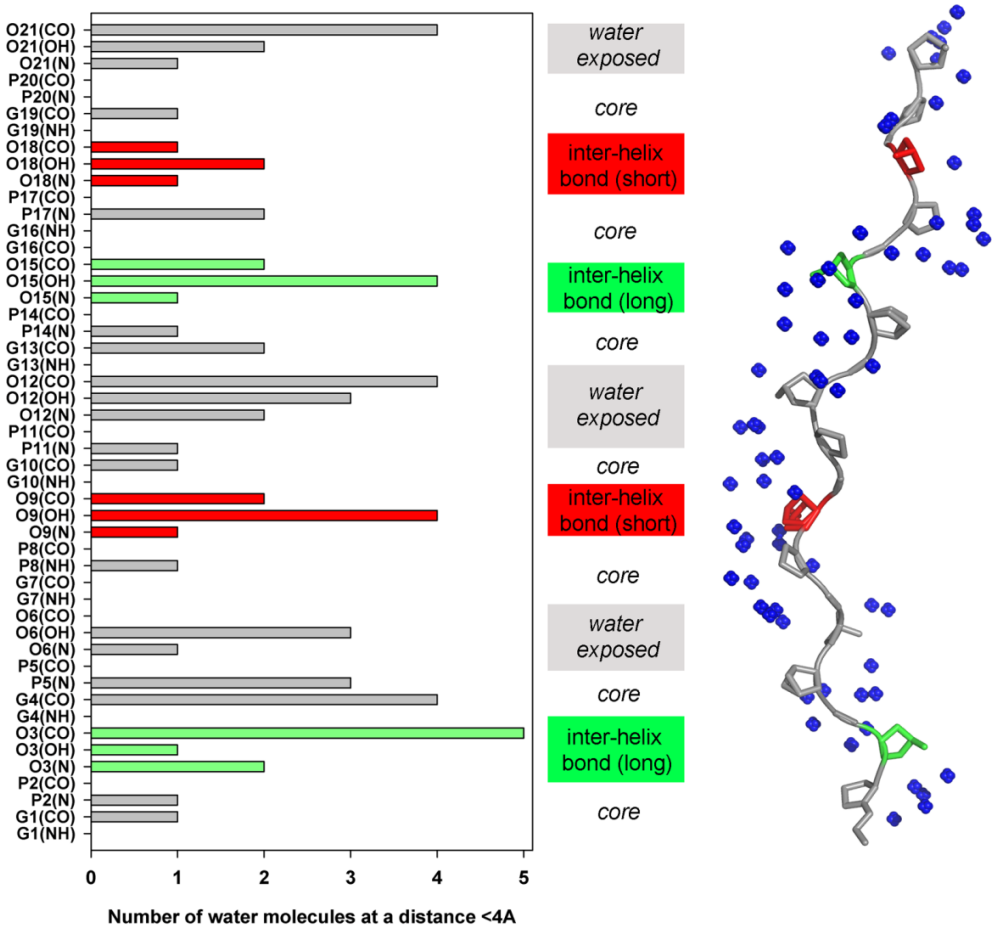


Figure 5.17 Number of water molecules in close vicinity with a single chain of 21 residues. The inter-helix long H bond has been shown in red and the inter-helix long bond has been done in green (expand). This demonstrates that the water shell is found mainly localised around the atoms of hydroxyproline residues, including the side-chain hydroxyl groups.

For both pairs of hydroxyproline residues, their hydroxyl groups also display short distances (<4 Å) with water molecules (Figure 5.17).

This vicinity supports water-mediated stabilization of the crystalline network, through hydroxyproline side chain hydration. The crystal structure displays 156 water molecules for one triple helix (52 per chain). This water shell is found to be mainly localised around the atoms of hydroxyproline residues, including the side-chain hydroxyl groups (Figure 5.17).

Along a chain of 21 residues, the hydration shell is helical, with clear alternating pattern of hydrated sites (mainly hydroxyproline residues) and hydrophobic cores (mainly glycine residues). This geometry of hydration sites corresponds to the structure of the triple helix, with glycine residues forming the core of the helix, and hydroxyproline residues on the outer surface.

Hydration of hydroxyproline side-chains has been previously proposed to play a role in triple helix stability [189, 190]. However, until today, this is the first time that water molecules are proposed to stabilise the crystalline pseudo-hexagonal network of triple helices.

5.7 Summary and findings

The 0.89 Å resolution structure of the $(GPO)_{10}$ collagen model peptide has allowed the examination of the atomic details of a model triple helix arranged in a crystalline pseudo-hexagonal network. Along with confirmation of generic features of model triple helices, such as abundant intra and inter-strand hydrogen bond networks within one triple helix, the localisation of glycine residues in the hydrophobic core of the triple stranded structure and the direct hydrogen bonding between triple helices via hydroxyproline side-chains, it was also shown that, proline rings can adopt two stable conformational states. The helical pitch was seen to vary around a 7/2 helix and that the stabilisation of the crystalline network has been water-mediated via hydroxyproline hydration, which bridged hydrogen bonding of helices.

Even though the successful racemic crystallisation of the smallest possible collagen model peptide of 16 residues, $(GPO)_5$-Gly was achieved and it diffracted at both 1 Å and 1.5 Å resolutions, the structure could not be solved and triple helical packing could not be understood. From solution studies on L-$(GPO)_5$-Gly, the melting point has been determined to be ~35 °C and isodichroic point has been found to be 200 nm. On heat-treating these peptides

at 70 °C for 8 minutes, decrease in size of the aggregates has been noticed, similar to other GPO peptides from previous chapter.

6 Effect of oxidation on triple helices and supramolecular assemblies

$(GOO)_{10}$ peptides have been characterized in this chapter to understand the differences occurring in the assembly properties of GPO and GOO collagen peptides. This will help explore possibilities of improving the food matrices.

The chapter has also investigated the impact of heating (cooking) on self-assembly of $(GPO)_7$ collagen model peptides. Tandem mass spectrometric analysis has been carried out to sequence the peptides and characterize the oxidative modifications occurring in amino acid sites in the peptide chain.

6.1 Characterization studies from literature

Mizuno et al. define the stability of triple helices by the melting point of Gly-Xaa-Yaa collagen model peptides. The amino acid positioning at the Xaa and Yaa position can regulate the MPs of the helices [191]. From previous studies, $(GPP)_{10}$ has been reported to have a higher melting temperature than the $(GOP)_{10}$ peptides [192]. Hydroxyproline is also known to be a very important stabilizing factor in collagen triple helices [193]. $(GPO)_{10}$ is the most stable and ideal collagen triple helix based on the positioning of hydroxyproline as Yaa in $(Gly\text{-}Xaa\text{-}Yaa)_n$ sequence [60]. Studies on $(GPO)_{10}$ suggest that its MP is ~ 68 °C [56].

Multiple studies on the crystal structures of $(GPO)_{10}$ collagen peptides have shown that factors like proline puckering, cis-trans stabilisations, torsion angles and stereo-electronic effects have an impact on the triple helical stabilities [69, 70, 114]. The proposed hypothesis has been investigated in this chapter which explains that higher stability of $(GOO)_{10}$ triple helix is because of a unique repeating intra-strand water mediated bridges linking Hyp in the Xaa position to the carbonyl oxygen of the 4(R)Hyp residue in the Yaa-1 position of the same strand. This intra-chain network has been discussed in the crystal structure of $(GOO)_{10}$ [69]. No solution studies have been conducted until today that support this intra self- stabilisation of the polyproline II helix.

6.2 Introduction

Redox proteomic approaches were taken up by Grosvenor et al., for proteomic analysis of heat induced amino acid modifications from the environment and from processing of protein

based skin, textiles and foods [8]. Such techniques of comprehensive amino acid damage profiling after in vitro digestion has also granted in depth insights into the nutritional quality of food [194]. Hydrothermal insult is also known to induce oxidative modification in proteins and peptides but the effect of this modification on food structure is poorly understood [8, 195].

In this chapter, extensive tandem mass spectrometric analysis has been performed to sequence the peptides and to characterise the oxidative modifications occurring in amino acid sites in the peptide chain, because of cooking. Changed assembly properties because of oxidation were recorded by AFM while studying $(GPO)_7$ peptides. Redox proteomic approaches have been used to profile the peptides and track modifications. A range of oxidations have been recognised in the middle parts of the $(GPO)_7$ peptide chain. The process by which oxidation takes place has been studied by examining the pre-solved structure of $(GPO)_7$.

6.3 Results and Discussion

6.3.1 Selection of $(GOO)_{10}$ peptides

The notable form of oxidative modification in relation to collagen is hydroxylation, especially proline. In the formative stages of collagen, the initial peptide sequence called α-chain undergoes modifications in the endoplasmic reticulum [31].

$(GPO)_{10}$ and $(GOO)_{10}$ peptides have been previously studied with CD and DSC and the higher stability of collagen triple helix of $(GOO)_{10}$ has been established [191]. In this chapter, the differences in the first order assembly formation by $(GPO)_{10}$ and modified peptide $(GOO)_{10}$ have been investigated with circular dichroism and differential scanning calorimetry. IR spectroscopy on nano-assemblies and crystals have been performed to study the conformational changes that could support $(GOO)_{10}$'s higher stability. To examine the structural differences in the supramolecular assembly formations of modified peptide $(GOO)_{10}$ from non-modified peptide $(GPO)_{10}$, atomic force microscopy and transmission electron microscopy have been used.

6.3.2 Differences in triple helix formations of $(GPO)_{10}$ and $(GOO)_{10}$ peptide crystals; Analysis by IR spectroscopy

Infrared Spectroscopy can be performed by measuring absorption, emission and reflection. This technique is used by chemists to determine the functional groups in molecules. In this chapter, infrared spectroscopy was used to recognize non-covalent bonds, polyproline helical factors, denaturation, hydrogen bonding etc..

For mounting the biological samples, in infrared microscopy, usually two kinds of windows are used (1) CaF_2 (Refractive index 1.4) and (2) Diamond window (Refractive index 2.37). The selection of window is based on the spectral range to be studied, chemical properties of sample and physical properties of the window. For this research, diamond windows were used because they were best suited to focus on the crystals and to obtain noise free spectra. Bruker V80v Fourier transforms infrared (FTIR) spectrometer and Hyperion 2000 IR microscope were used to reach a high signal-to-noise ratios. The experiment was carried out at Australian synchrotron, in collaboration with Dr Celine Valery.

6.3.2.1 Sample preparation

The unheated crystal samples were prepared by the usual crystallization trials described in chapter 5. The heat treated crystals, analysed in section 6.3.2.3 were prepared as follows; the peptide solutions to be crystallized were subjected to heating at 70 °C for 8 minutes and then cooled down to room temperature 20 °C. Crystallization of these heat-treated peptide solutions was carried out by hanging drop vapour diffusion method. The hanging droplets were placed on cover slips along with reservoir solution, in equal ratios. These cover slips were inverted and placed on reservoir solutions. Crystals appeared after 2-3 weeks of incubation time.

After crystallization, the fresh hanging droplet was taken from the cover slip of the crystal tray and encapsulated in the diamond window and mounted on the microscope stage as shown in Figure 6.1C and Figure 6.1D. Figure 6.1A shows the $(GOO)_{10}$ crystals obtained at conditions of 7 mg/ml incubated in 4 °C for 2-3 weeks in a 96 well (2 sub-well) plate. Figure 6.1B shows the $(GPO)_{10}$ crystals obtained at 4 mg/ml concentration at a reservoir composition of 22% PEG 400. The water subtracted raw infrared spectra obtained from the unheated and heated crystal samples of $(GPO)_{10}$ and $(GOO)_{10}$ are shown in Figure 6.A.

Figure 6.B shows the 2nd derivative spectra on the unheated and heated samples of $(GPO)_{10}$ and $(GOO)_{10}$.

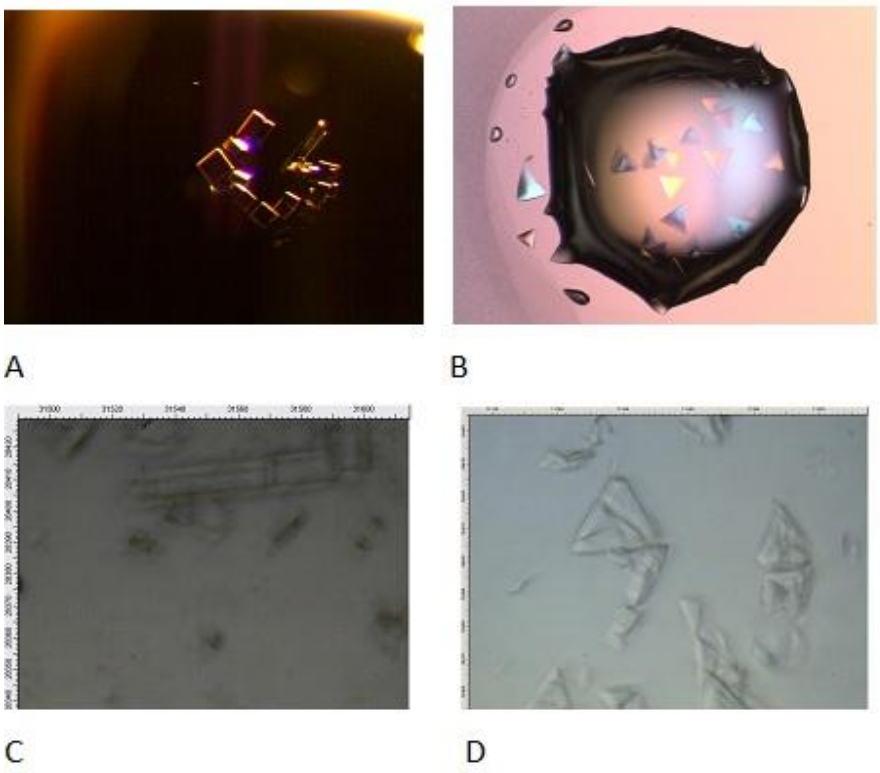


Figure 6.1 (A) $(GOO)_{10}$ crystals at concentration of 7 mg/ml PEG 400 (B) $(GPO)_{10}$ at 4 mg/ml concentration PEG 400 (C) $(GOO)_{10}$ crystals encapsulated in diamond window (D) $(GPO)_{10}$ encapsulated in diamond window.

6.3.2.2 Infrared spectroscopic analysis on crystals samples

Crystal samples were used because samples for infrared spectroscopy usually require concentrated samples preferably in the form of assemblies, instead of dilute peptide solutions to obtain clear infrared spectra. Since crystals are formed as a result of high solute concentration during the vapor diffusion process of crystallization, these proved to be the ideal samples.

Prystupa et al. have shown that the bands at 1645–1657 cm^{-1} are characteristics of random coils and 1660 cm^{-1} are characteristic of triple helices [144, 147, 153]. In some studies of triple helical collagen and collagen like peptides, the absorption bands for amide I appear at approximately 1650 cm^{-1} and amide II peaks appear at around 1560 cm^{-1} [147].

Three peaks for amide I region of polyproline II helix have been reported at 1650 cm^{-1}, 1640 cm^{-1} and a peak of maximum absorbance at 1632 cm^{-1} [146]. The shift of amide I, II and III peaks to low wavenumbers, broadening of amide I peak and increase in intensity of amide III peak are associated with increased intermolecular interactions (hydrogen bonding) in collagen [148].

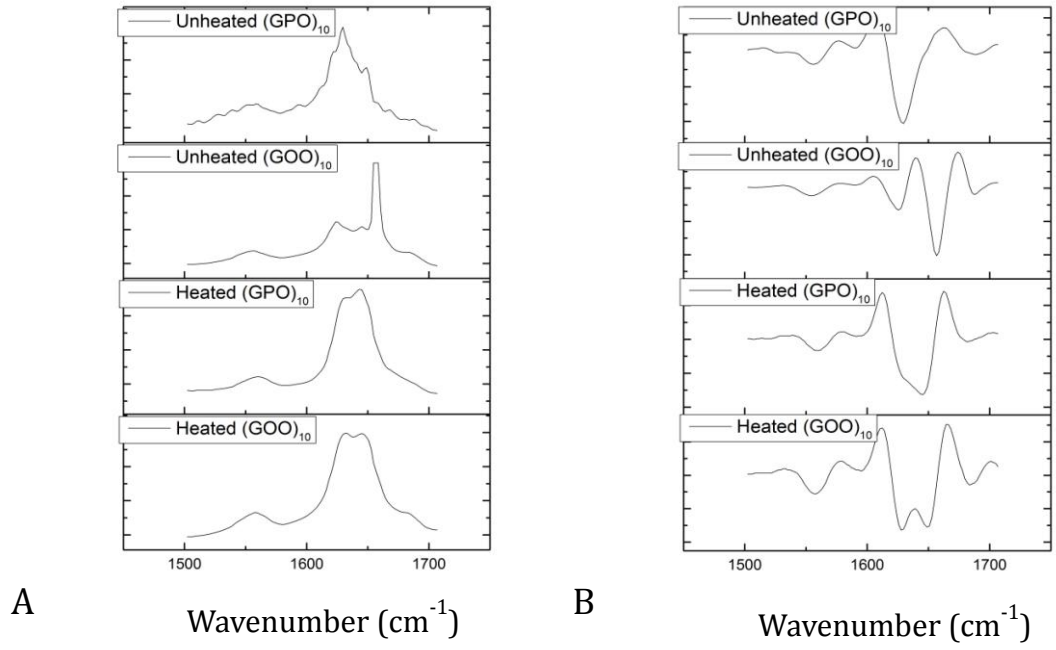


Figure 6.2 IR Absorbance vs. Wavenumber (A) Raw spectral peaks for unheated and heated $(GOO)_{10}$ and $(GPO)_{10}$ crystal samples (B) Second derivative spectra for unheated and heated $(GOO)_{10}$ and $(GPO)_{10}$ crystals

6.3.2.3 GPO and GOO crystals

From this analysis on $(GPO)_{10}$ crystal samples, formed from the unheated peptide solutions, H-bonding in helices was confirmed from the presence of bands at 1629 cm^{-1}. This peak had shifted to higher wavenumbers, at around ~ 1648 cm^{-1} for the crystals formed from the heat-treated peptide solutions. Peak shifting to higher wavenumbers represent the breakage of H-bonds in collagen (peaks at 1648 cm^{-1} symbolize random coils). Thus uncoiling of the $(GPO)_{10}$ helix has been inferred for the heat treated samples.

For $(GOO)_{10}$ crystals formed from unheated peptide solutions, the amide absorption peak was observed at 1650 cm-1. This is characteristic of a collagen triple helix. Crystals formed from the heat treated peptide solutions showed no wavenumber shifting and the peaks

remained intact. The heated helical assemblies of $(GOO)_{10}$ is believed to have intact H-bonds, as compared to $(GPO)_{10}$ crystals after the same temperature treatments. Stable H-bonds have indicated higher stability for the $(GOO)_{10}$ peptides as compared to the $(GPO)_{10}$ peptides. A peak at ~ 1632 cm^{-1} has also been seen for heated $(GOO)_{10}$ crystals, indicating presence of the polyproline II helix (Table 6.1).

Table 6.1 Infrared spectroscopy on unheated and heated crystals of $(GPO)_{10}$ and $(GOO)_{10}$

CRYSTALS	Raw Infrared Spectra (Water subtracted)			Second derivative Spectra		
	Amide I (cm^{-1})	Amide I Shoulder (cm^{-1})	Amide II (cm^{-1})	Amide I (cm^{-1})	Amide I Shoulder (cm^{-1})	Amide II (cm^{-1})
Un-Heated **$(GPO)_{10}$**	1629	1648	-	1629	1648	-
Un-Heated **$(GOO)_{10}$**	1651	1642	-	1651	-	1540
Heated **$(GPO)_{10}$**	1645	-	1540	1648	-	1545
Heated **$(GOO)_{10}$**	1650	-	1540	1650 1632	1629	1540

6.3.3 Triple helix formations for $(GPO)_{10}$ and $(GOO)_{10}$ peptides; Analysis by CD

Samples were prepared at concentration of 1 mg/ml at an incubation time of 1 hour at 4 °C. The buffers used were water and 20 mM PBS for the CD studies [48]. The MRE vs. wavelength scans were set up from 260 nm to 180 nm. The scan was carried out at 20 °C. In Figure 6.3A shows that the $(GPO)_{10}$ and $(GOO)_{10}$ peptides show triple helix formation. The maximum at 225 nm indicates the presence of triple helices [82, 94].

6.3.3.1 Stability of triple helices

The stability of a peptide is measured by the temperature point until which triple helix stays intact and does not melt [196]. Beyond this the inter and intra helical hydrogen bonds break.

Fallas et al. have reported the melting point of $(GPO)_{10}$ as ~60-65 °C [179]. The MP for $(GOO)_{10}$ has been reported to be 80.5 °C by CD studies and 81.8 °C by DSC studies [191]. Frank et al. have reported the melting point of $(GOO)_{10}$ as ~ 76.6 °C by CD and 74.3 °C by DSC studies [191, 196].

For both $(GPO)_{10}$ and $(GOO)_{10}$ peptides, apart from stability, the differences in triple helix formations, winding, re-winding and analysis of melting points and re-naturation points have been studied in details with circular dichroism (Figure 6.3).

6.3.3.2 Stability studies - Analysis of melting point, re-naturation point, unwinding and re-winding of triple helices with CD

Figure 6.3 shows the melting of helices in five steps. The black and red curves are used to define $(GOO)_{10}$ and $(GPO)_{10}$ peptides respectively.

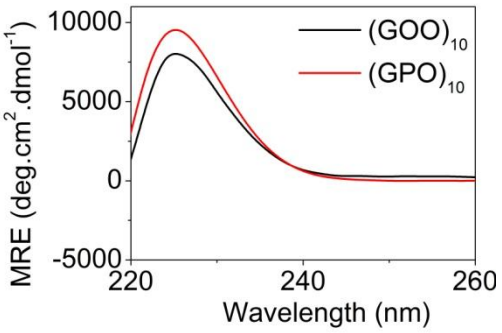
Step 1 and Figure 6.3A, confirms the presence of triple helices with a maximum at 225 nm, which depicts the presence of triple helix [82, 94]. Helices have been confirmed for both $(GPO)_{10}$ and $(GOO)_{10}$ peptides.

In step 2 and Figure 6.3B, unwinding of triple helices is shown. Temperature scans were performed from 20 °C to 90 °C, at consistent scanning rates of 1°C/minute. The data pitch was kept constant at 1°C. This was carried out at a wavelength of 225 nm. A sigmoid curve was seen for both $(GOO)_{10}$ and $(GPO)_{10}$ peptides. The melting point of $(GPO)_{10}$ peptides is found to be between 60 and 65 °C and the MP of $(GOO)_{10}$ peptides is found to be between 70 and 75 °C.

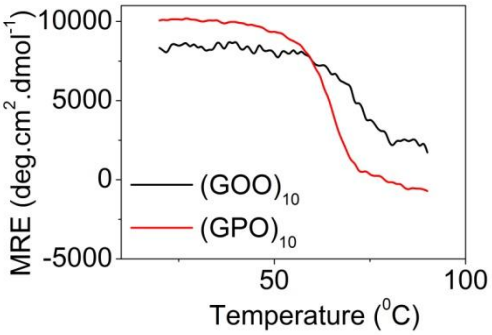
In step 3, Figure 6.3C, MRE vs. wavelengths scan, at 90°C was performed. A maximum at 225 nm is observed for $(GOO)_{10}$.

As per the data, triple helices were still present for $(GOO)_{10}$ peptides. However, for $(GPO)_{10}$ peptides, no maximum was observed. This could have been brought about by the

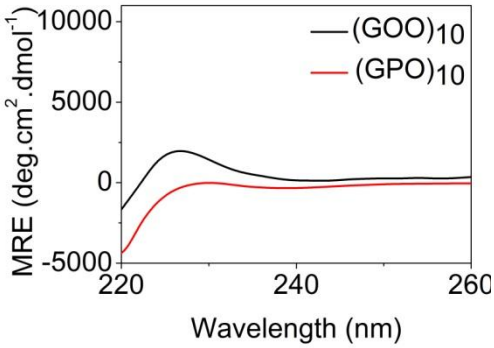
unwinding of the triple helices or breaking of hydrogen bonds. It is speculated that the extra stability of $(GOO)_{10}$ peptides is due to its intrachain water bridge.



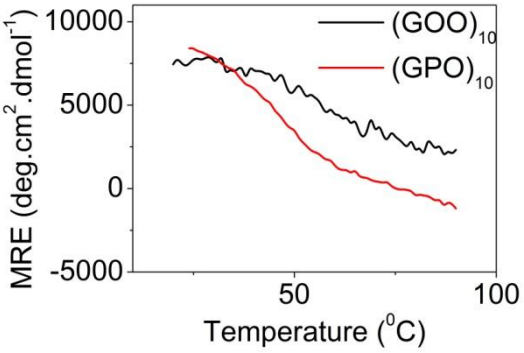
A



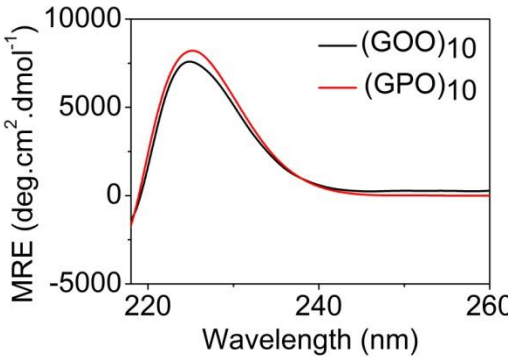
B



C



D



E

Figure 6.3 Unwinding and rewinding of triple helices; black curve depicts $(GOO)_{10}$ and red curve is $(GPO)_{10}$; (A) Wavelength scan at 20 °C (B) Temperature scan at 225 nm, from 20 °C to 90 °C (C) Wavelength scan at 90 °C (D) Temperature scan at 225 nm, from 90 °C to 20 °C (E) Wavelength scan back

Previously published crystal structure of $(GOO)_{10}$ by Schumacher et al. in 2005 also explains that the polyproline II helix conformation of the unfolded peptide chains, individually stabilises the intra-stand water mediated bridge, linking the Hyp in Xaa position to carbonyl oxygen of the 4(R) Hyp residue in the Yaa-1 position of the same strand [69].

In step 4 (Figure 6.3D), the temperature scans from 90 °C to 20 °C at 225 nm have been shown.

In step 5, Figure 6.3E shows the wavelength scan from 260 nm to 180 nm once again at 20°C in which the triple helices have re-appeared again, for both $(GOO)_{10}$ and $(GPO)_{10.}$ Mizuno et al. have reported for $(GPO)_{10}$ and $(GOO)_{10}$ peptides, that the MP of acetyl-$(GOO)_{10}$-NH_2 was found to be ~ 80 °C and MP for acetyl-$(GPO)_{10}$-NH_2 was found to be ~ 74 °C [191]. From studies of Kar et al., the MP of $(POG)_{10}$ has been found to be 60 °C. In this chapter, the higher stability of GOO over GPO has been confirmed and the melting point of $(GPO)_{10}$ has been recorded as 65 °C and MP of $(GOO)_{10}$ was recorded as 70 °C.

In 2005, studies by Schumacher reasoned why the CD signal at 225 nm stayed positive for the denatured state of $(GOO)_{10}$ [191] and they hypothesized that a polyproline II-like conformation could be responsible [69]. This hypothesis was supported by the studies of Brahmachari et al [197] where they suggested that the poly 4(R) hydroxyproline formed a highly stable hydrogen bonded left handed helix in aqueous solution whose CD spectra had a negative at 205 nm and a positive peak at 225 nm.

6.3.4 Thermodynamics of triple helix unwinding of $(GPO)_{10}$ and $(GOO)_{10}$ peptides; Analysis by DSC

Differential scanning calorimetry (DSC) has been used to analyze the thermal transitions and enthalpy changes. DSC data analysis has been done using the 2 state scaled, trimer to monomer model. The software used for data analysis is Nanoanalyse. It uses the Van't Hoff equation given below

$$\frac{dlnK}{dT} = \Delta HvH/(RT^2)$$

Equation 6.1

The thermodynamic expression is

$$\frac{d}{dt}(\ln K) = \Delta H/(RT^2)$$

Equation 6.2

K = Equilibrium constant of process

T= absolute temperature

vH= Van't Hoff Enthalpy

ΔH=Enthalpy

6.3.4.1 Sample preparation

Samples were prepared as per procedures from Kar et al [48]. $(GPO)_{10}$ and $(GOO)_{10}$ peptides were prepared at 5 mg/ml concentrations, in 20 mM PBS buffer (150mM NaCl) and were incubated for 2 days at 4 °C.

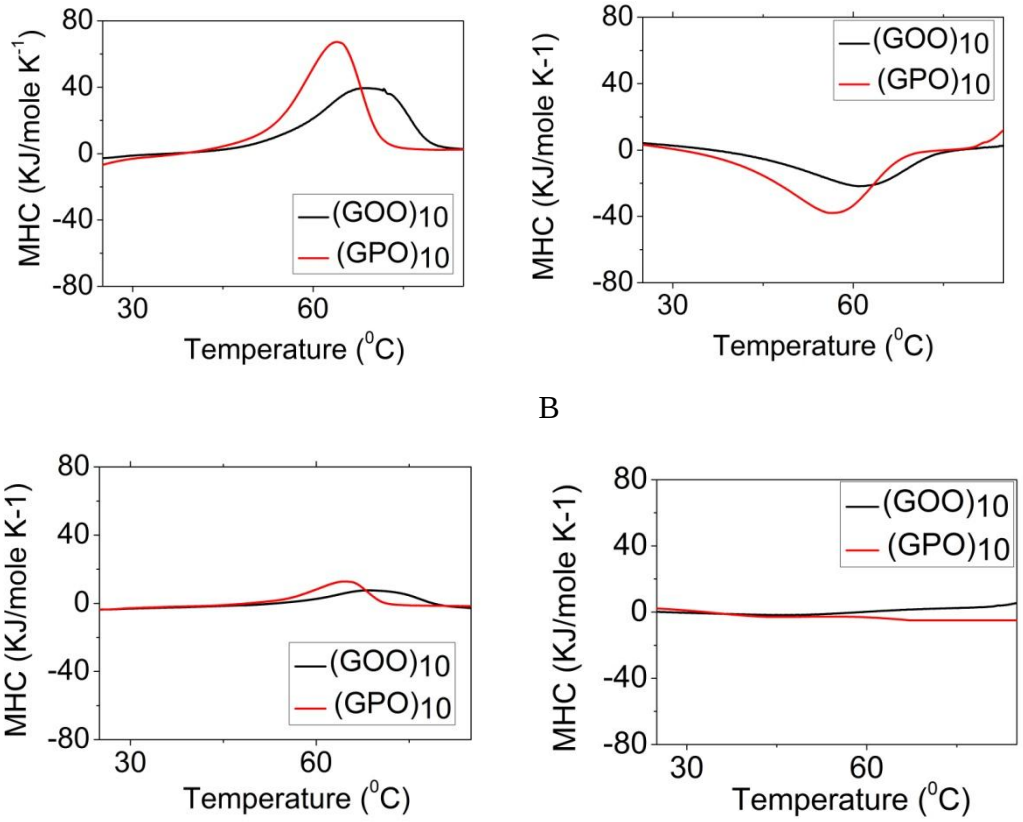


Figure 6.4 Differential scanning calorimetric (melting point) analysis of $(GPO)_{10}$ and $(GOO)_{10}$ peptides. Scanning speed was 1 ^{0}C/minute. Buffer used is 20 mM PBS; $(GPO)_{10}$ is denoted by black and $(GOO)_{10}$ is denoted by red curve (a) Heating at 5 mg/ml concentration (b) Cooling at 5 mg/ml concentration (c) Heating at 1 mg/ml concentration (d) Cooling at 1 mg/ml concentration

Heating (Figure 6.4A) and cooling (Figure 6.4B) scans were performed at 1°C/minute, to understand the conformational changes with thermal transitions. The cell volume was 600 µl.

6.3.4.2 Data analysis

The raw data has been converted into molar heat capacity with the unit KJ/(mol K^{-1}) with the software Nanoanalyse. Molecular weight of peptides, $(GOO)_{10}$ is 2850.85 Daltons and $(GPO)_{10}$ is Daltons. The melting points and enthalpies have been determined by using the two state scaled model.

6.3.4.3 Interpretation of DSC curves

From Figure 6.4A, at 5 mg/ml concentration, $(GOO)_{10}$ has a melting point of 70 °C and the enthalpy is 288.92 KJ/mol with an error factor (A_w) of 0.11 (Error factor less than 0.5 indicates that data is believable [198]). For $(GPO)_{10}$ peptides, at a concentration of 5 mg/ml, the melting point is 63.75 °C and enthalpy change was calculated to be 350.93 KJ/mol.

The enthalpy change was seen to be smaller for $(GOO)_{10}$ peptides as compared to $(GPO)_{10}$. The thermal transitions were also broader for $(GOO)_{10}$ as compared to $(GPO)_{10}$. From experimental data, it can be said that $(GOO)_{10}$ is more stable as compared to $(GPO)_{10}$. The renaturation point is ~ 66 °C for $(GOO)_{10}$ and for $(GPO)_{10,}$ the re-naturation point is ~ 59 °C, at concentrations of 5 mg/ml, see Figure 6.4B. Broad thermal transitions, smaller enthalpy changes and higher transition temperatures for $(GOO)_{10}$ have also been previously reported in literature [191]. Mizuno et al have also shown that DSC showed broader thermal transitions for $(GOO)_{10}$ as compared to the $(GPO)_{10}$ peptides [191]. With $(GOO)_{10}$, no complete conformation changes were observed when the helix transformed into a monomeric non-helical state at high temperatures. Thus, $(GOO)_{10}$ was more stable than $(GPP)_{10}$ and $(GPO)_{10}$ peptides.

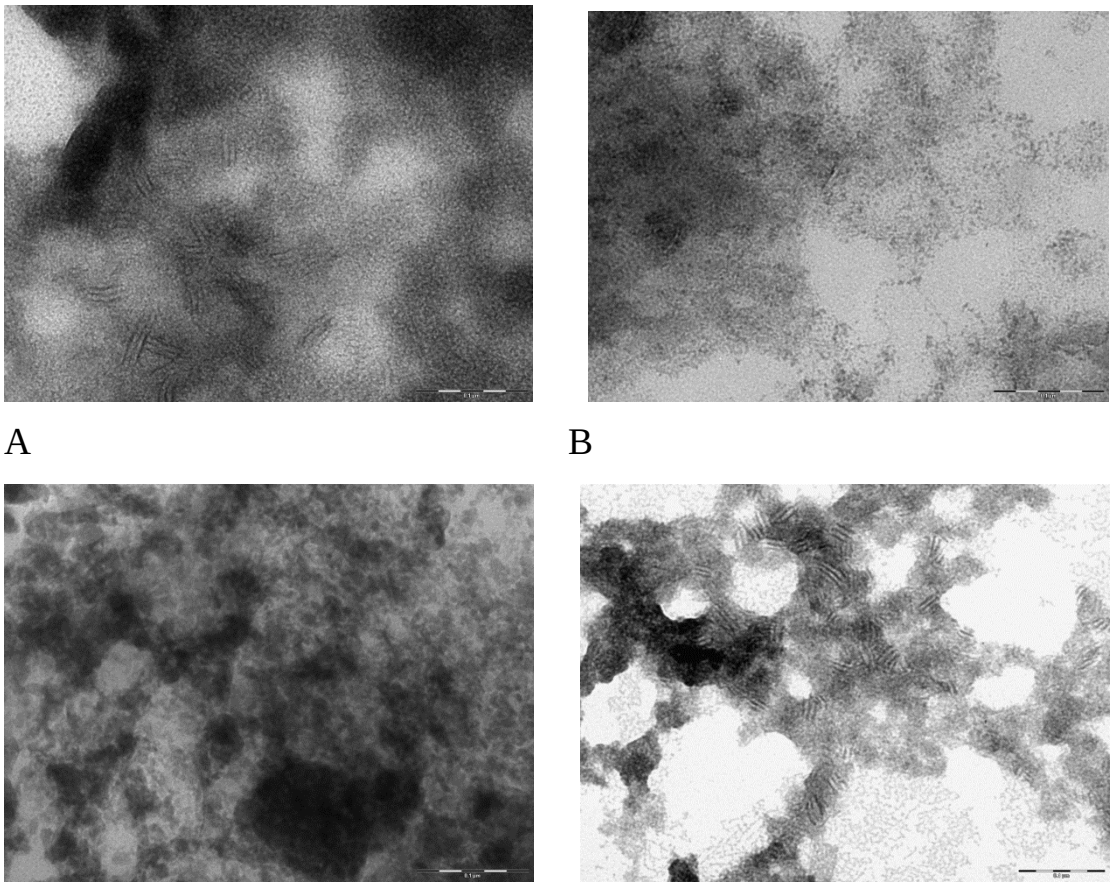
Figure 6.4C is the DSC analysis at 1 mg/ml concentration of peptides. The melting point for $(GOO)_{10}$ peptides is seen at ~ 69 °C, while the enthalpy is 378.27 KJ/mol. For $(GPO)_{10}$ peptides, the melting point is ~ 70 °C and enthalpy is around 377.31 KJ/mol. The error factor (A_w) for $(GPO)_{10}$ analysis was 0.06. This slight difference in results for $(GPO)_{10}$ at different concentrations could have been because of insufficient sensitivity. The renaturation point for 1 mg/ml concentration of peptides could not be successfully determined (Figure 6.4D). It is speculated that the triple helices did not fully wind back at low concentrations on cooling, at the same scan rate.

6.3.5 Nano-scale assembly formations for $(GPO)_{10}$ and $(GOO)_{10}$ peptides; Analysis by TEM

6.3.5.1 Sample preparation

Samples of concentrations of 7 mg/ml of $(GPO)_{10}$ and $(GOO)_{10}$ peptides were prepared as per procedure described elsewhere in the thesis. These were heat treated and subsequently cooled down to 20 °C. Heating was performed at incubation temperatures of 37 °C, 50 °C, 60 °C and 70 °C for 8 minutes.

6.3.5.2 TEM analysis



A B C D

Figure 6.5 $(GOO)_{10}$ analysis by TEM; Peptides were heat treated at specific temperatures for 8 minutes and were cooled down till room temperature (20 °C). All figures are at 0.1 micron scale. (A) incubation at 37 °C for 8 minutes showed possible aggregates; some artefacts were also visible due to carbon in the TEM grids (B) incubation at 50 °C for 8 minutes was dominated by artefacts (C) incubation at 60 °C for 8 minutes showed non-conclusive species (D) branched filamentous assemblies were seen at 70 °C, on incubation for 8 minutes (some artefacts due to TEM grid carbon were also visible). TEM artefacts were differentiated from supramolecular assemblies based on following references [131], [132].

For $(GOO)_{10}$ peptides (Figure 6.5) incubated at 37 °C for 8 minutes, aggregates were seen. Some artefacts were also visible due to the presence of carbon in the TEM grids. The sample incubated at 50 °C for 8 minutes is dominated by artefacts. The sample incubated at 60 °C for 8 minutes formed non-conclusive species of $(GOO)_{10}$ peptides. Branched filamentous assemblies were also seen in the sample incubated at 70°C for 8 minutes (some artefacts due to TEM grid carbon also showed up).

For $(GPO)_{10}$ peptides (Figure 6.6), the sample at 37 °C incubation for 8 minutes and a separate sample at 50 °C incubation for 8 minutes show non-conclusive species. The formation of branched filamentous assemblies was observed, when $(GPO)_{10}$ peptides were incubated at 60 °C for 8 minutes.

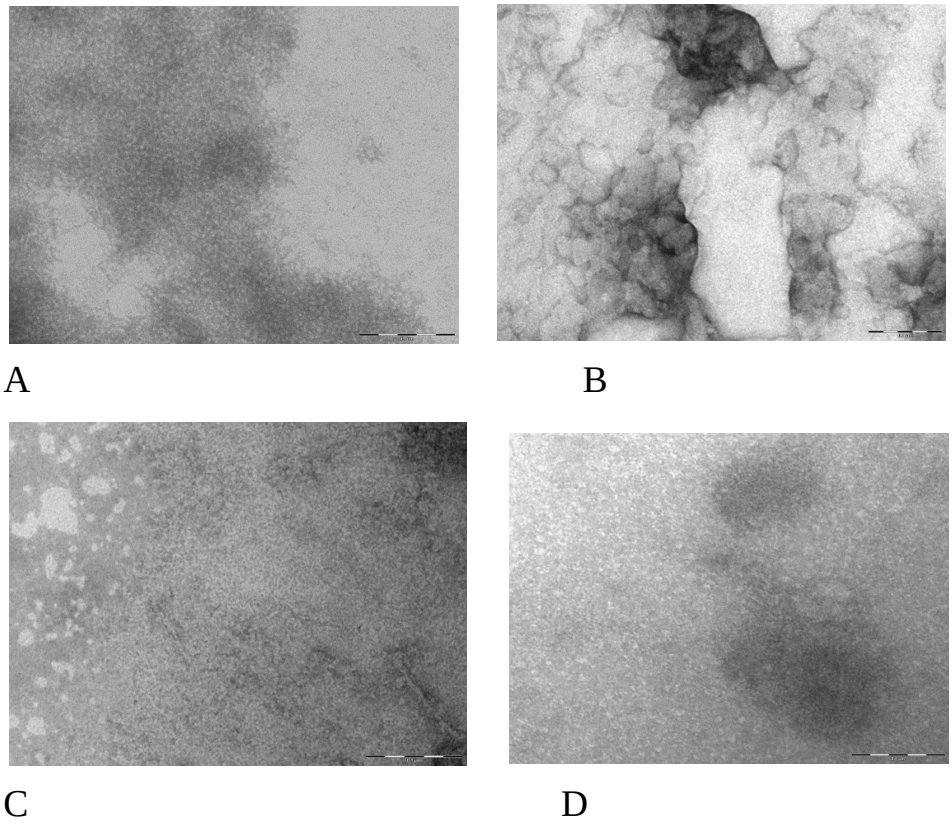


Figure 6.6 $(GPO)_{10}$ analysis by TEM; All figures are at 0.1 micron scale. $(GPO)_{10}$ peptides were heat treated at specific temperatures for 8 minutes and cooled down till room temperature and (A) incubation at 37 °C for 8 minutes formed non-conclusive species (B) incubation at 50 °C for 8 minutes formed non-conclusive species (C) incubation at 60 °C for 8 minutes initiated formation of branched filamentous assemblies which were similar to ones seen by Kar et al. (D) Possible supramolecular assemblies were seen at 70 °C, on incubation for 8 minutes. TEM artefacts were differentiated from supramolecular assemblies based on following references [131], [132].

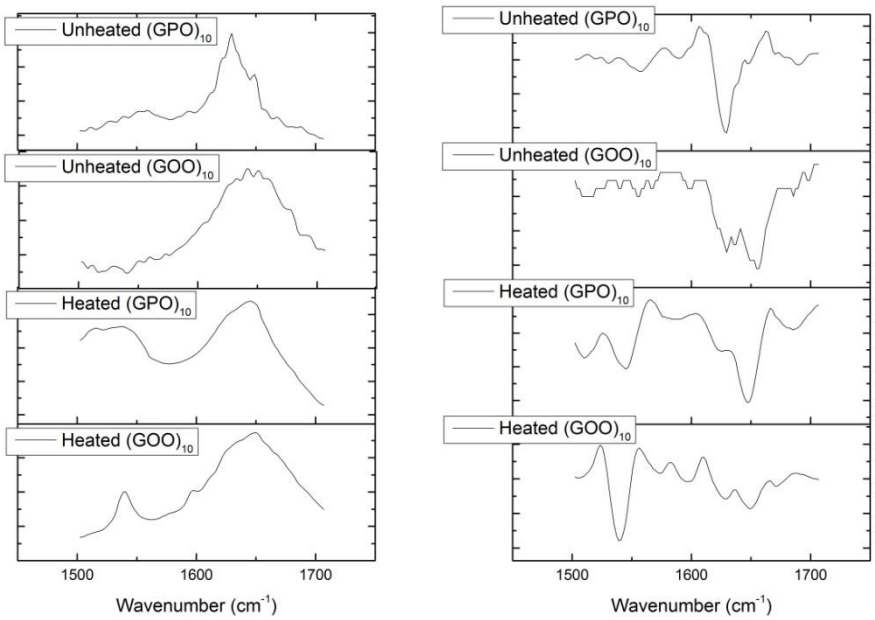
Kar et al. had seen the similar kind of assemblies for $(GPO)_{10}$ peptides [48]. Possible supramolecular assembly formations were noticed at 70 °C after 8 minutes of incubation.

6.3.6 Infrared analysis on nano-scale assemblies

Table 6.2 Infrared spectroscopy on supramolecular assembly (unheated and heated)

SUPRAMOLECULAR ASSEMBLIES	Raw Infrared Spectra (Water subtracted)			Second derivative Spectra		
	Amide I (cm^{-1})	Amide I Shoulder (cm^{-1})	Amide II (cm^{-1})	Amide I (cm^{-1})	Amide I Shoulder (cm^{-1})	Amide II (cm^{-1})
Un-Heated **$(GPO)_{10}$**	1629	1648	1559	1629	-	1556
Un-Heated **$(GOO)_{10}$**	1658	1625	1556	1657	1625	1559
Heated **$(GPO)_{10}$**	1644	-	1560	1645	-	1560
Heated **$(GOO)_{10}$**	1632	1645	1558	1657	-	1657

The amide I (1600-1690 cm^{-1}, C=O stretching), amide II (1480-1575 cm^{-1}, CN stretching, NH bending) and amide III (1229 -1301 cm^{-1}, CN stretching, NH bending) mostly reveal the conformational changes of the proteins and peptides [145, 199]. The bands at 1645–1657 cm^{-1} are termed as random coils. Peaks at 1660 cm^{-1} denote the triple helices [144]. The bands at 1690 cm^{-1} show the presence of helices of aggregated collagen like peptides [145]. Infrared spectroscopy has been performed to analyze the properties of the aforementioned assemblies. The assembly samples for analysis are shown in Figure 6.5D and Figure 6.6D.



A B

Figure 6.7 IR absorbance vs. Wavenumber (cm^{-1}) (A) Raw IR spectrum for unheated and heated $(GOO)_{10}$ and $(GPO)_{10}$ supramolecular assemblies (B) Second derivative spectrum for unheated and heated $(GOO)_{10}$ and $(GPO)_{10}$ supramolecular assemblies.

6.3.6.1 $(GPO)_{10}$ and $(GOO)_{10}$ peptides

The reference unheated $(GOO)_{10}$ and $(GPO)_{10}$ peptide samples at 7 mg/ml in 20 mM PBS, pH-7 were also analysed by IR spectroscopy, along with the heat-treated samples (at 70 °C for 8 minutes).

Peaks at 1660 cm^{-1} are characteristics for triple helices and 1650 cm^{-1} are characteristics of triple helical H-bonding. The 2nd derivative spectra for $(GPO)_{10}$ and $(GOO)_{10}$ peptides shown in Figure 6.7 and Table 6.2 shows that H-bonding is prominent in $(GPO)_{10}$ helical assemblies. After heat treatment, random coils have been noticed at 1645 cm^{-1} for $(GPO)_{10}$ peptides.

For $(GOO)_{10}$ peptides, helices at ~1658 cm^{-1} were observed for the reference unheated assemblies. After heating, these helices stay intact at 1657 cm^{-1}. This supports the previous results that $(GOO)_{10}$ assemblies are more stable than assemblies of $(GPO)_{10}$.

6.3.7 Effect of pH on triple helix formation of $(GPO)_{10}$ and $(GOO)_{10}$ peptides; Analysis by DSC

6.3.7.1 DSC analysis

Differential scanning calorimetry was performed to study the melting point and enthalpy for triple helices. pHs were chosen at three ranges, pH 3.5, pH 7 and pH 8.0. Those pH values were used to meet the range of pK_a of peptide terminals and to check the assembly's response to various pH conditions.

Since NH_2-$(GPO)_n$-COOH and NH_2-$(GOO)_n$-COOH were used, selections were carried out based on the pK_a of terminals. For NH_2-glycine, the pK_a is 9.6 and the pK_a of hydroxyproline-COOH is 1.92. The buffers used for various pHs were (1) 10 mM citrate buffer at pH 3.5 (mol weight = 294.10 gm/mol) (2) pH 7- sodium phosphate buffer 10 mM (mol weight of Na_2HPO_4 141.96gm/mol, $NaH_2PO_4.1H_2O$ =137.99 gm/mol); and (3) pH 8, Tris 10mM buffer (mol wt 121.1 gm gm/mol).

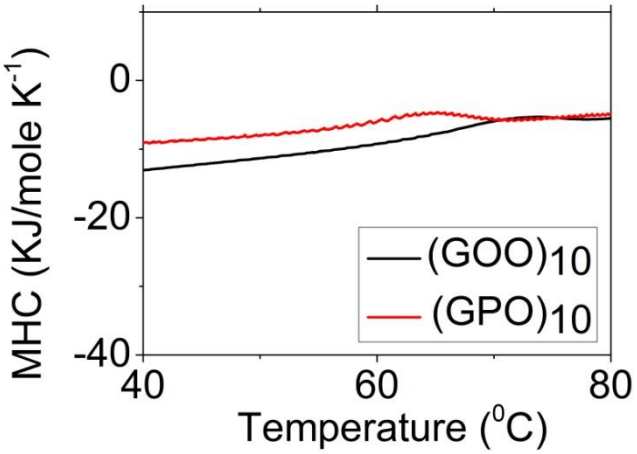


Figure 6.8 DSC analysis on $(GPO)_{10}$ and $(GOO)_{10}$ peptides, at concentration of 1 mg/ml, at pH 3.5.

6.3.7.2 DSC analysis at pH- 3.5

From Figure 6.8, the melting point of $(GOO)_{10}$ peptides was 71.21 ° C at a concentration of 1 mg/ml at pH 3.5, with an error factor of 0.034. This is nearly same as the melting point of $(GOO)_{10}$ peptides at pH 7. Enthalpy was 393.219 KJ/mol.

For $(GPO)_{10}$ peptides, DSC analysis at pH 3.5 showed melting point at 63.7 °C and enthalpy was 450.477 KJ/mol with an error (A_w) factor of 0.026.

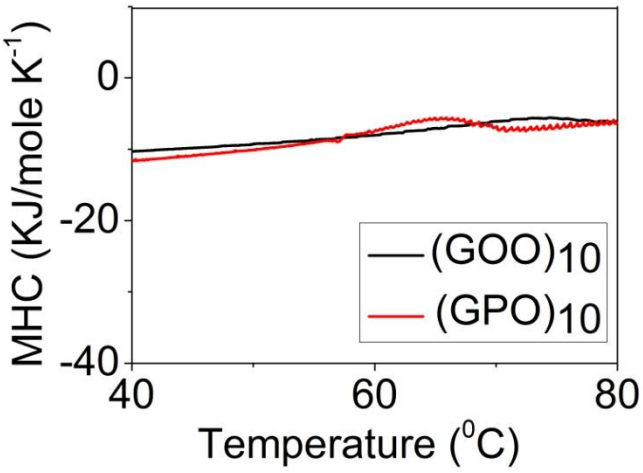


Figure 6.9 DSC on $(GPO)_{10}$ and $(GOO)_{10}$ peptides at a concentration of 1 mg/ml and at pH 8.0

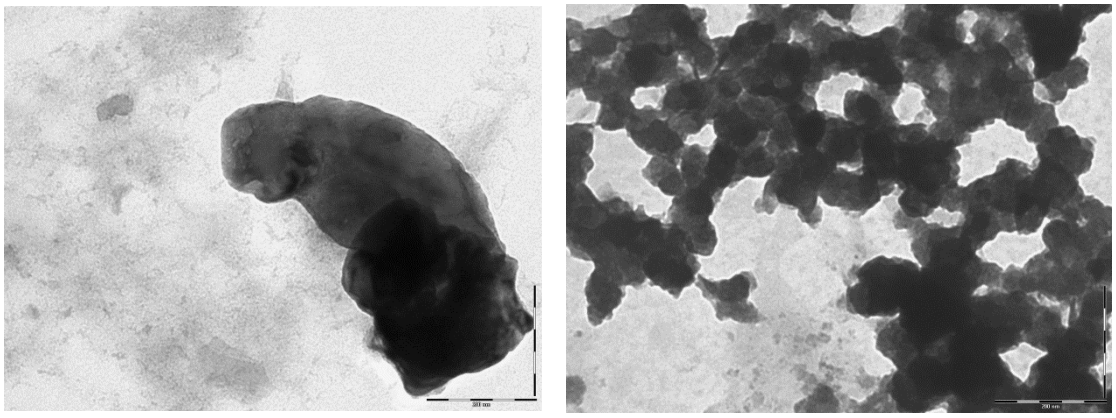
6.3.7.3 DSC analysis at pH- 8.0

From Figure 6.9, the MP for $(GOO)_{10}$ peptides at pH 8.0 was 72.30 °C. Enthalpy was found to be 422.26 KJ/mol, with an error factor of 0.017. The software Nano-Analyse was used.

For $(GPO)_{10}$ peptides at 1 mg/ml and pH 8, melting point was found to be 64.90 °C, an enthalpy was 476.85 KJ/mol, with an error factor of 0.02.

6.3.8 Effect of pH on nano-scale assembly formation of $(GOO)_{10}$ and $(GPO)_{10}$ peptides; Analysis by TEM

Nano-scale structure formations were analysed at a peptide concentration of 7 mg/ml as per procedures from Kar et al. and they were modified [48].



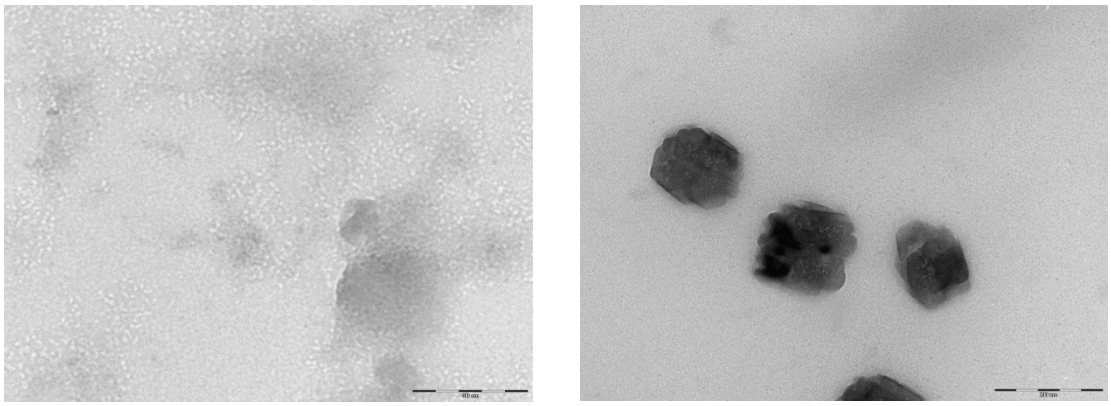
A B

Figure 6.10 TEM analysis at pH 3.5 (A) $(GPO)_{10}$ peptides at 7 mg/ml at 200 nm; No clear assemblies were seen, grids had overloading and looked like artefacts. Images were non-conclusive. (B) $(GOO)_{10}$ peptides at 7mg/ml at 200 nm; aggregations were seen. TEM artefacts were differentiated from supramolecular assemblies based on following references [131], [132].

Incubation time of 1 hour and room temperature of 20 °C was chosen. Three ranges of pHs were chosen which were, pH 3.5, pH 7 and pH 8.0. These were based on the pK_a of the peptide terminals.

At pH 3.5, TEM images of both $(GPO)_{10}$ peptides (Figure 6.10A) and $(GOO)_{10}$ peptides (Figure 6.10B), seem to have overloading and images are non-conclusive.

At pH 8, with an incubation time of 1 hour, at room temperature (Figure 6.11A), tiny nano-scale assemblies were observed for $(GPO)_{10}$. $(GOO)_{10}$ shows no assemblies, except for a few salt crystals (Figure 6.11B). The pH favouring supramolecular assembly formations could not be determined.



A B

Figure 6.11 TEM analysis at pH- 8.0 (a) $(GPO)_{10}$ peptides at 7mg/ml at 200nm; very small aggregates were seen (b) $(GOO)_{10}$ peptides at 7mg/ml at 100nm; salt crystals was seen and no assemblies were visualized. TEM artefacts were differentiated from supramolecular assemblies based on following references [131], [132].

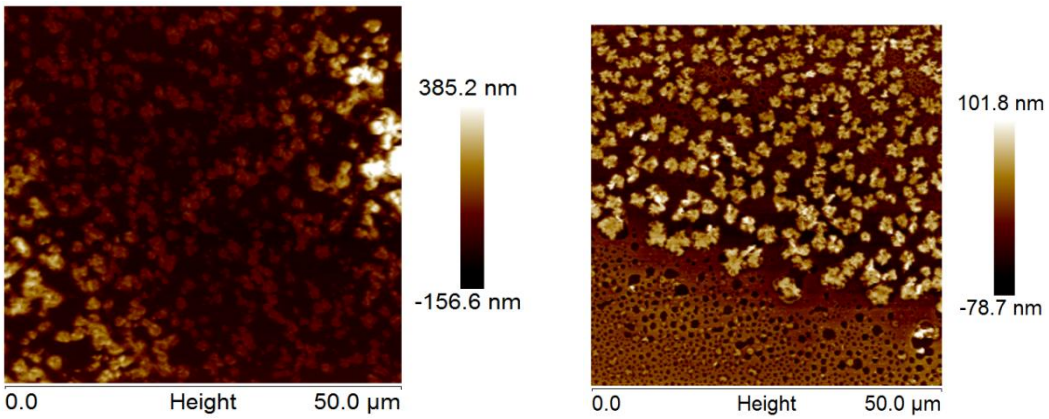
6.3.9 Cooking of collagen mimetic peptides

In earlier chapters, it has been observed that oxidation events within the protein primary structure result in changes in self-association properties of protein and its derivative peptides [15]. This may prove to be of importance in the manufacture of reconstituted foods from meat proteins. Based on this observation, the collagen model peptides have been cooked or been introduced to hydrothermal insult, in the following sections.

Following it, a before and after cooking comparative study of the peptides was performed, to see the changes in their assembly properties.

6.3.9.1 Effect of cooking on $(GPO)_7$ peptides

$(GPO)_7$ was cooked at 90 °C for 1 hour. Supramolecular assemblies were observed by atomic force microscopy after cooking. Figure 6.12A shows the $(GPO)_7$ peptides in their unheated forms, in 20 mM PBS solution. When the same peptides at 7 mg/ml were cooked at 90 °C for 1 hour and then cooled down (Figure 6.12B), nano-assemblies with some changes in their morphologies and decreased sizes were observed. The supramolecular assemblies of heated $(GPO)_7$ peptides had reduced heights of ~101 nm from the original heights ~ 385.2 nm for the unheated assemblies. Scan size of 50 microns and the scan angle of 90 degrees were used. Average samples/lines were fixed at 256 at a scan rate of 0.100 Hz.



A B

Figure 6.12 AFM imaging of $(GPO)_7$ self assembled structures before and after cooking the peptides at 90 °C for 1 hour at 7 mg/ml concentration in 20 mM PBS buffer. Noticeable decrease in height for heat-treated assemblies were recorded by AFM. AFM parameters, fixed tip velocity, scan angle and samples per line ratio and scan size were kept the same for both samples (A) unheated control samples of before cooking show convoluted structures of heights ~ 385.2 nm (B) heat treated samples of after cooking show convoluted structures of decreased heights ~ 101.8 nm. AFM artefacts were differentiated from supramolecular assemblies based on the following reference, [132].

A comparative study was also been performed by atomic force microscopy on collagen fibrils. Because of the longer filamentous structure of collagen, to facilitate imaging, the scan size was fixed at 30 microns. A reference buffer scan was initially performed as shown in Figure 6.13D. A reference unheated scan for collagen was also attempted, but due

to collagen's non soluble nature, the imaging was not successful. Fibrils were observed in collagen proteins that were subjected to heating (cooking) in Figure 6.13B. These were also observed with TEM as shown in Figure 6.13A.

To analyze the differences in supramolecular assemblies of $(GPO)_7$ collagen model peptides, mass spectrometric analysis has been performed on both the cooked and un-cooked samples and oxidations were recognized.

6.3.9.2 Calculation of % oxidation in the cooked and uncooked $(GPO)_7$ peptide supramolecular assemblies

Mass spectrometry was used to examine oxidations in both the uncooked and cooked $(GPO)_7$ peptides. AFM samples from above studies have been subjected to MALDI-TOF. Three sets of samples were prepared for the analysis.

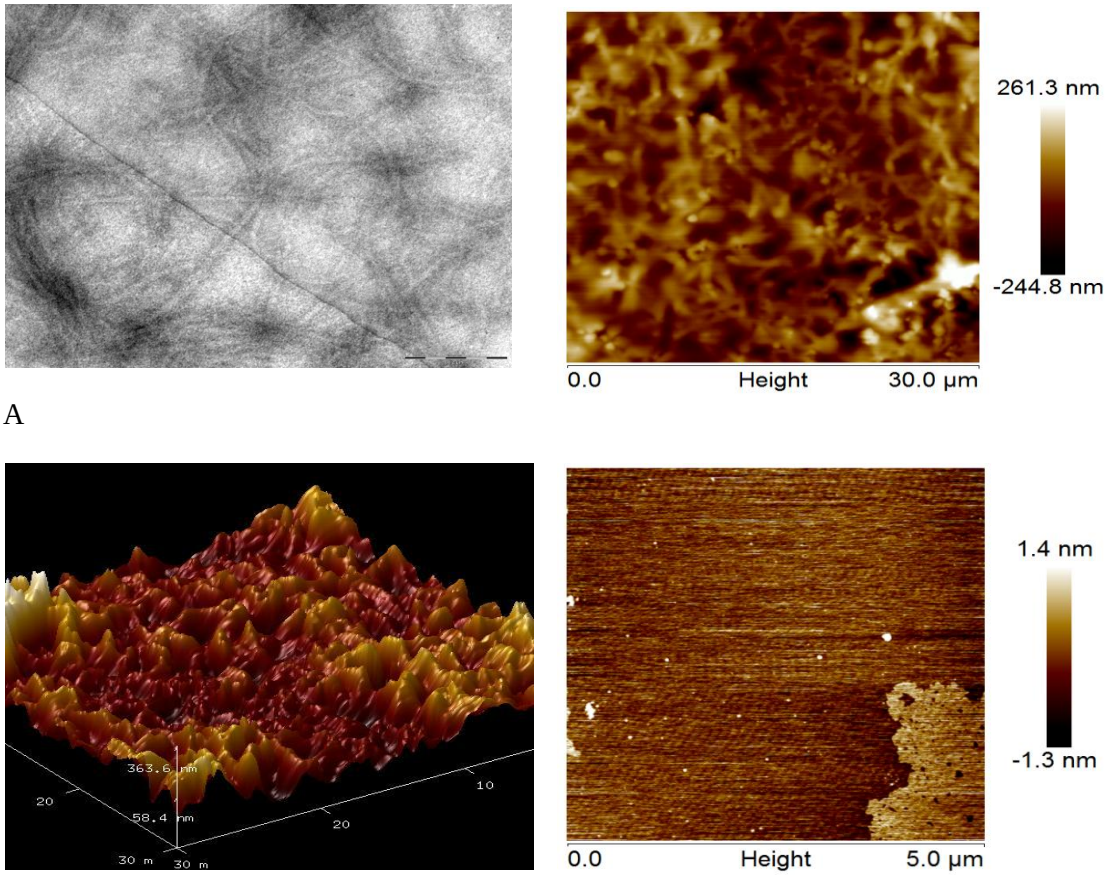


Figure 6.13 Collagen type I fibrils at 1 hour incubation at 90 °C at a concentration of 7 mg/ml in 20 mM PBS buffer (A) TEM imaging at 200 nm (B) AFM imaging at 30 µm scan size (c) AFM 3D profiling of the collagen sample (d) AFM imaging of reference 20mM PBS buffer.

6.3.9.3 Sample preparation

The 1st set of collagen sample was prepared in 20 mM PBS buffer at pH 7 and incubated for 1 hour at 90 °C at a concentration of 7 mg/ml. The 2nd set of sample was prepared in the same conditions but in 10 % acetic acid and the 3rd set of sample was prepared in water at the same conditions. These set of buffers were chosen for analysis of % oxidations because they have been previously observed to support supramolecular assembly formations after heating.

6.3.9.4 Data analysis

Protocols were used as described in Dyer et al [78]. The rationale behind performing this experiment was to examine the peptide solution with the maximum amount of oxidation occuring in it. After oxidation, an oxidation % would be calculated for all the peptide solutions, the one with most oxidation % would be further subjected to MSMS to recognize sites of oxidations.

For evaluation of the oxidation %, basic mass spectrometric analysis was performed on various peptide solutions and the m/z base peak for peptide, in X-axis was picked in mass spectrum. For $(GPO)_7$ samples, the peak was observed at 1889 Da. After locating the base peak, peaks with oxidation adducts were looked for. These peaks are the ones with additional masses of +32 and +16 on the base peaks. The oxidation % was calculated by dividing the peak area of base peak from the peak area of oxidized peak, for all sample solutions.

From Figure 6.14, the % of oxidation has been shown on the uncooked and cooked samples on the di-sodiated base peaks. Oxidations peaks were highest for the cooked samples and least for the uncooked ones. For peptides in water, the oxidation percentages were higher for both samples in comparison to the oxidation percentages in PBS buffer. It is speculated that this could be because of the presence of salt in PBS buffer samples that does not get properly removed, even after de-salting of the samples during thin layer affinity HCCA anchor chip preparations. It is usual for the unheated material to have a little oxidation, while undergoing sample preparation, treatment and during the introduction of MALDI target plate into the mass spectrometer.

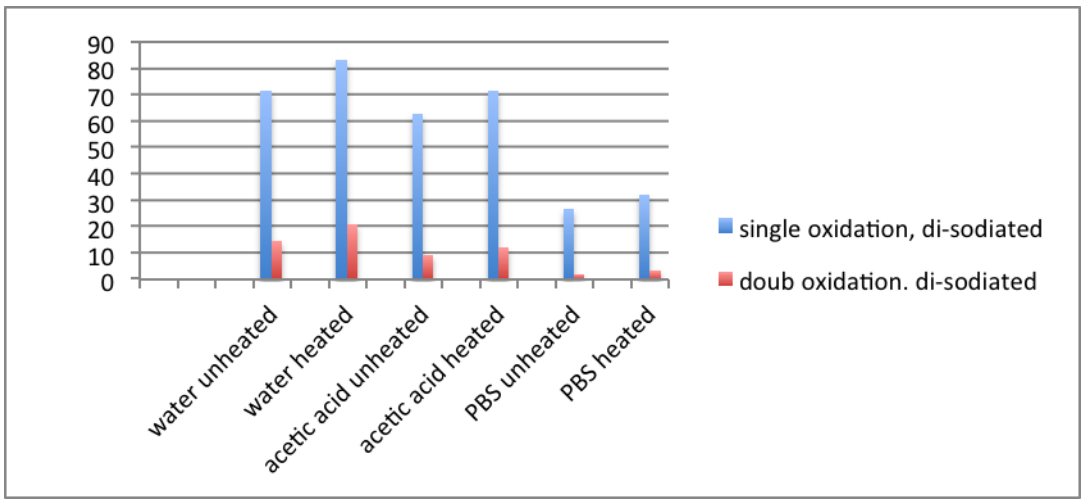


Figure 6.14 Calculation of "Oxidation % for $(GPO)_7$ peptides vs. peptide solutions in various buffers". MS analysis of $(GPO)_7$ peptides in three different buffers (water, 10% acetic acid and PBS buffer), to calculate oxidation % has been shown in the graph. Profiled oxidations from PBS buffer samples have been further examined with MSMS that are shown in Figure 6.15 and in Table 6.3, Table 6.4 and Table 6.5

Tandem mass spectrometry was carried out on the $(GPO)_7$ heated PBS samples among the heated samples to study the oxidized residues. Fragmentation of these oxidized peaks shown in Figure 6.14, yield the y and b ions. Based on the differences occurring in the sequence of y and b ions of the oxidized peak, from the sequence of y and b ions of the base peak, the specific oxidized residues have been recognized. These have been discussed in details in the next section on tandem mass spectrometry.

6.3.9.5 Tandem Mass spectrometry

In the MS spectra of $(GPO)_7$ peptides, having mass/charge as their X-axis, base peak of 1932, singly oxidized peak of 1948 and the doubly oxidized peak of 1964 were seen (Figure 6.15). These were subjected to further fragmentation for an MSMS.

The details of entire row of the di-sodiated y and b ions of the peaks of masses 1932, 1948 and 1965 for $(GPO)_7$ peptides have been summarized in the tables below. From the y and b ion fragmentation profiles of the three di-sodiated peaks (base peak, singly oxidized peak and doubly oxidized peak), the oxidized residues for $(GPO)_7$ peptides were located. 44 was added to all the y and b ions of the base peak to account for sodiation as shown in Table 6.3. Table 6.4 shows the fragmentation profile for singly oxidized peak and Table 6.5 shows the fragmentation profile for doubly oxidized peak.

Table 6.3 –y and b ions of base peak fragmentation (m/z- 1932)

		Fragmentation $(GPO)_7$		di-sodiated (+44)	di-sodiated (+44)
		B IONS	Y IONS	B IONS	Y IONS
b1	y21	58	-	102	-
b2	y20	155	1831.8	199	1875.8
b3	y19	268	1734.7	312	1778.7
b4	y18	325	1621.74	369	1665
b5	y17	422	1564.72	466	1608.7
b6	y16	535	1467.67	579	1511.6
b7	y15	592	1354.62	636	1398.6
b8	y14	689	1297.606	733	1341.6
b9	y13	802	1200.553	846	1244.5
b10	y12	859	1087.505	903	1131.5
b11	y11	956	1030.484	1000	1074.4
b12	y10	1069	933.431	1113	977.43
b13	y9	1126	820.384	1170	864.3
b14	y8	1223	763.362	1267	807
b15	y7	1336	666.309	1380	710.309
b16	y6	1393	553.262	1437	597.26
b17	y5	1490	496.24	1534	540.24
b18	y4	1603	399.187	1647	443.18
b19	y3	1660	286.14	1704	330.14
b20	y2	1757	229.118	1801	273.118
b21	y1	-	132.066	1918	176.06

Table 6.4 y and b ions of singly oxidized peak fragmentation; m/z of (1948)

Singly oxidized (1948)			
Fragmentation $(GPO)_7$		SINGLY OXIDATED (+16)	SINGLY OXIDATED (+16)
B IONS	Y IONS	B IONS	Y IONS
b1	y21	118	-
b2	y20	215	1891.8
b3	y19	328	1794.7
b4	y18	385	1681
b5	y17	482	1624.7
b6	y16	595	1527.6
b7	y15	652	1414.6
b8	y14	749	1357.6
b9	y13	862	1260.5
b10	y12	919	1147.5
b11	y11	1016	1090.4
b12	y10	1129	993.43
b13	y9	1186	880.38
b14	y8	1283	823
b15	y7	1396	726.3
b16	y6	1453	613.26
b17	y5	1550	556.24
b18	y4	1663	459.18
b19	y3	1720	346.14
b20	y2	1817	289.118
b21	y1	1934	192.06

Table 6.5 y and b ions of doubly oxidised peak fragmentation m/z of (1964)

		Doubly oxidized (1964)	
Fragmentation $(GPO)_7$		DOUBLY OXIDATED (+32)	DOUBLY OXIDATED (+32)
B IONS	Y IONS	B IONS	Y IONS
b1	y21	134	-
b2	y20	231	1907.8
b3	y19	344	1810.7
b4	y18	401	1697
b5	y17	498	1640.7
b6	y16	611	1543.6
b7	y15	668	1430.6
b8	y14	765	1373.6
b9	y13	878	1276.5
b10	y12	935	1163.5
b11	y11	1032	1106.9
b12	y10	1145	1009.43
b13	y9	1202	836.3
b14	y8	1299	839
b15	y7	1412	742.309
b16	y6	1469	629.26
b17	y5	1566	572.24
b18	y4	1679	475.18
b19	y3	1736	362
b20	y2	1833	305
b21	y1	1950	208.06

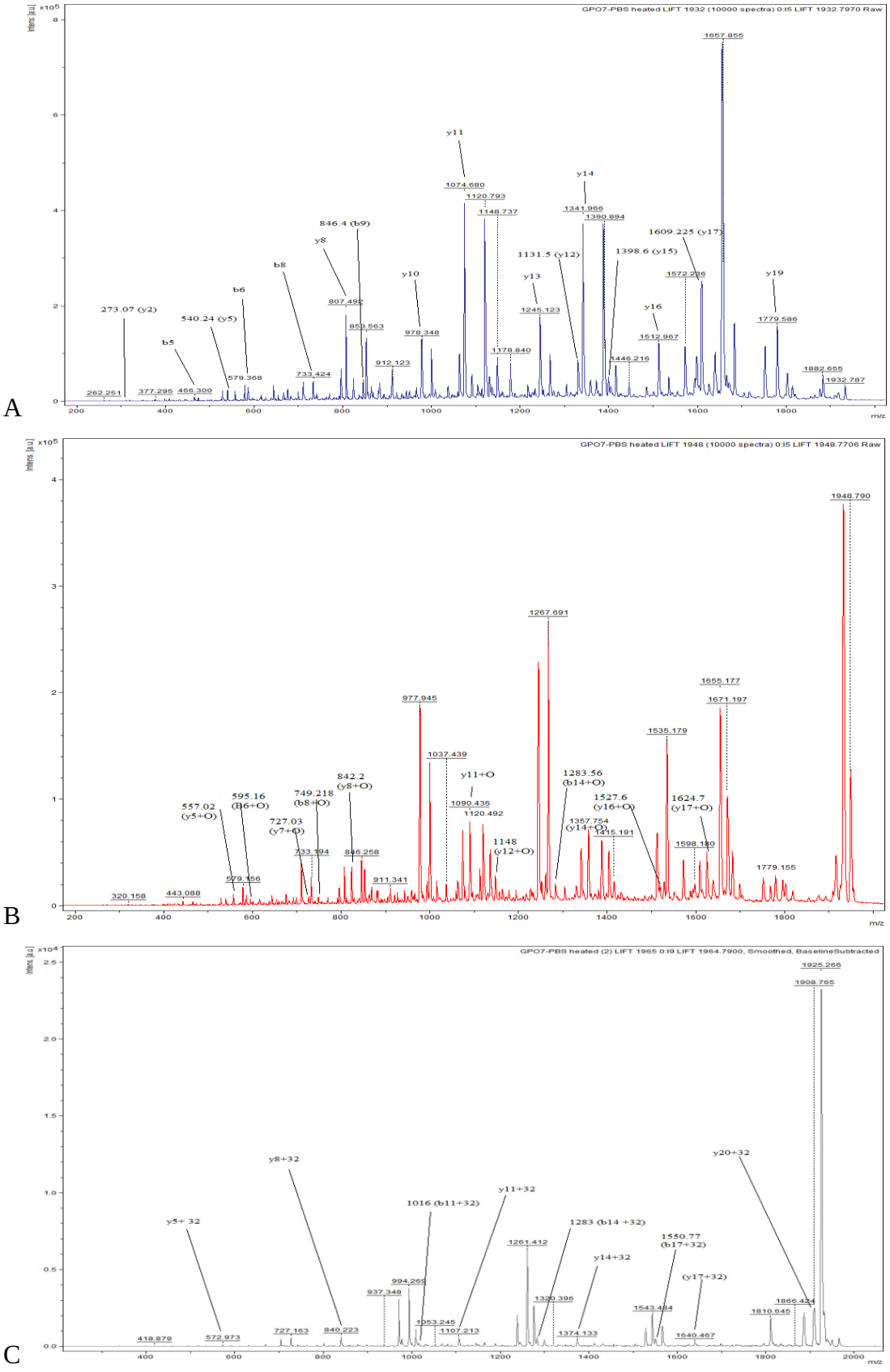


Figure 6.15 (A) Fragmentation profile of 1932, showing the y and b ions (B) Fragmentation profile of 1948, showing the y and b ions (C) Fragmentation profile of 1965 showing y and b ions.

From the fragmentation profile of single oxidation, it was observed that the 3rd, 4th and 5th prolines and the 2nd hydroxyproline were oxidized. From the fragmentation profile of double oxidation, it was observed that the 3rd, 4th, 5th and 6th prolines were oxidized and the 3rd, 4th and 5th hydroxyprolines were oxidized too.

The staggered arrangement of the helices is believed to be the primary reason behind the mid-strand zones getting targeted by hydrothermal insult (Figure 6.16).

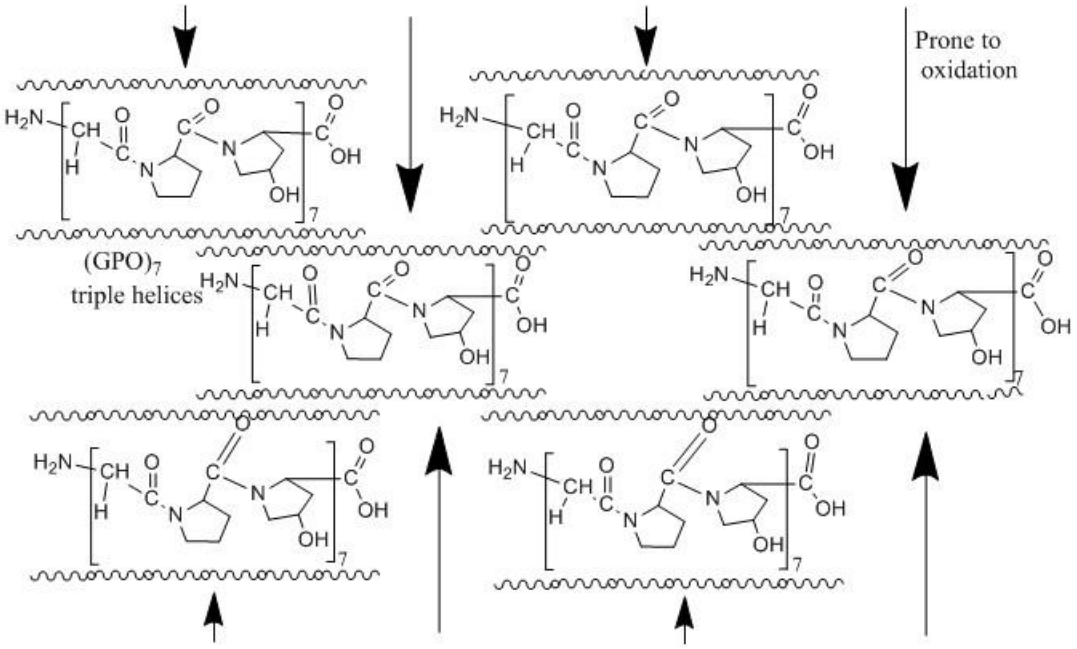


Figure 6.16 Chemical structure of $(GPO)_7$ triple helices in a staggered arrangement. Sites that are prone to oxidation are exposed because of staggering.

6.4 Discussion and findings

Differences in the first order assembly, triple helix formations and the higher order assemblies, nanoscale structure formations of $(GPO)_{10}$ and $(GOO)_{10}$ have been examined in the first part of this chapter. $(GOO)_{10}$ peptides proved to be more stable than the $(GPO)_{10}$ peptides by CD and DSC studies. Mizuno et al [191] have suggested 3 possible reasons for this higher stability, i.e.

(1) steric hindrances imposed by the proline ring

(2) the inter-chain hydrogen bonds between GlyNH…OC(Xaa) of adjacent chains and

(3) hydroxylation of proline residues in Yaa position to 4 (R)-hydroxyproline, such that either hydroxyl group of 4R Hyp forms water bridges between backbone groups or the inductive effect of hydroxyl group influences the ring puckering of pyrrolidine ring and changes the cis/trans ratio of peptide bond.

Other reasons for the additional stability could be puckering. It has been observed that the up-puckered conformation of 4R-hydroxyproline in the Xaa position was different to the down-puckered conformation of the Xaa position in $(GPO)_{10}$ peptides, which could lead to additional stability. But, even after most of the 4 (R) Hyp residues in the Xaa position took up the up-puckered conformation in $(GOO)_{10}$, there was not much difference in the dihedral angles of this structure and the other GPP and GPO triple helical structures. Moreover, the super helix symmetry of $(GOO)_{10}$ was also found to be 7/2, which was similar to other helical symmetries.

Based on the aforementioned observations, it has been speculated that the most favoured decisive factor, for the $(GOO)_{10}$ peptide to be more stable than $(GPO)_{10}$, can be the intra-strand stabilization of polyproline II helix. In this intra-chain water network, the hydroxyl moieties of the Xaa 4(R) Hyp residues hydrogen bind to the water molecule, which is also hydrogen bonded to the carbonyl oxygen of the Yaa 4(R) Hyp residue, located one triplet NH2-terminal to the Xaa 4 (R) Hyp. However, there has not been any literature on solution studies supporting this intra-chain water network way of stabilization.

This chapter provides further evidence on the intrastrand water stabilization of polyproline II like structure with solution studies via infrared spectroscopy on crystals. Polyproline II like structural amide I peaks have been seen to remain stable even after heat treatment on crystals. The infrared spectroscopy on heated assemblies have also supported the higher stability of $(GOO)_{10}$ peptides. As reported in other studies, our CD results also show a stable peak at 225 nm even after denaturation, supporting the presence of a stable polyproline II like structure in each unfolded monomer strand. A broader transition temperature for $(GOO)_{10}$ has been noticed in DSC, along with a higher melting point in comparison to $(GPO)_{10}$.

Changes in nano-scale assemblies due to cooking have been observed by AFM. Tandem mass spectrometric analysis has been carried out to sequence the peptides and

characterize the oxidative modifications due to cooking. The middle parts of the $(GPO)_7$ peptide sequence were found to be highly prone to oxidation. The staggered arrangement of the helices is believed to be the primary reason behind peptide's mid-strand zones getting targeted by hydrothermal insult.

7 Summary and conclusions

7.1 Introduction

The objectives of this thesis, as outlined in chapter 1 were

(i) to investigate the effect of GPO chain length on self-assembly.

(ii) to crystallize and solve the (GPO) collagen model peptides of varying chain lengths and discover the shortest possible (GPO) peptide to undergo crystallization.

(iii) to study the differing assembly properties because of oxidation; $(GPO)_{10}$ and $(GOO)_{10}$ peptides

(iv) to understand the effect of hydrothermal insult on collagen peptide chain and its impact on the supramolecular assemblies.

7.2 Characterization of triple helices

7.2.1 Effect of GPO chain length on GPO triple helix formation

From CD characterization studies of GPO peptides, the triple helix formation was found to be directly proportional to chain length, concentration of mixture and time of incubation. Longer chain lengths of peptides were seen to exhibit higher helical stabilities. The most supportive pH for triple helix formation was pH 7, which is in accordance with previous studies [48]. pH 3.5 favored triple helical stability the most, after pH 7.

7.2.2 Effect of chain length on GPO helical triple helical stability

Helical stability was found to be directly proportional to chain length. Triple helical reversibility of collagen peptides has also been confirmed in previous studies [138]. Triple helical reversibility was re-confirmed for all the chosen GPO peptide systems in our experiments. Chain length was found to be inversely proportional to critical concentration of helical melting, following DSC experiments.

7.2.3 $(GPO)_7$ and amidated $(GPO)_7$ stability studies

Amidation of end terminals was seen to increase the stability of the $(GPO)_7$ helices; Amidated $(GPO)_7$ peptides also have a higher melting temperature as compared to $(GPO)_7$ peptides, as determined from CD melting studies.

7.2.4 High ionic strength environment support triple helical assembly

In previous studies by Gurry et al., effect of salt was found significant in supporting the collagen triple helical formations and stability [200].

High salt environment strongly supports assembly formations by GPO peptides. At high salt conditions (213 mM NaCl), the triple helical content for the GPO peptides was seen to be higher than the helical content in normal salt environments (150 mM NaCl). In high salt environment, CD cooperative transition was supported by both $(GPO)_7$ and $(GPO)_{10}$ peptides, whereas in usual salt conditions, cooperative transition was only supported by the $(GPO)_{10}$ peptides.

7.2.5 Manual melting analysis of GPO triple helices

Manual melting studies with circular dichroism indicated that with $(GPO)_7$ and $(GPO)_{10}$ peptides, the minima associated with random coils shifted to higher wavelengths with temperature increments. It was speculated that this could have resulted from breaking of H-bonds.

The minima shift (dynamic coils) of ~ 200 nm to higher wavelengths occurred more in $(GPO)_{10}$ than in $(GPO)_7$ peptides. It is speculated that this was because of prevalence of more H-bonds in $(GPO)_{10}$.

7.3 Characterization of supramolecular assemblies

7.3.1 Self-association kinetics of GPO peptides

Turbidity studies on $(POG)_{10}$ have recorded the formation of self-associated structures, arising from 1st stage assembly of triple helices [48]. $(POG)_7$ had no indication of supramolecular assemblies at 7 mg/ml at any temperature. $(POG)_8$ had a small rise in turbidity (supramolecular assembly) at 48 °C at 7 mg/ml. $(POG)_{10}$ recorded self-associations at various temperatures, while $(POG)_{12}$ had high propensity of aggregation even at low concentrations and low temperatures, than seen for $(POG)_{10}$ [48].

In studies of Kar et.al, on $(POG)_{10}$ peptides, there was no aggregation even after incubation for 1 week at high concentrations of 14 mg/ml at 1 °C - 4 °C [133]. In this research work it was shown that $(GPO)_{10}$ achieved supramolecular assemblies at high salt environment, in 100 mM PBS buffer (213mM NaCl) at 1% concentration (1mg in 100 ul) at

4 °C and at 1 week incubation time. $(GPO)_{10}$ also formed assemblies at 37 °C, at 1% concentration when incubated for a week. Peptides of longer chain lengths like $(GPO)_7$ and $(GPO)_{10}$ achieved triple helical assemblies, along with supramolecular assemblies right from low concentrations (0.2 mg.ml), unlike shorter peptides like $(GPO)_5$ and $(GPO)_3$.

7.3.2 Analysis by small angle X-ray scattering (SAXS)

From previous studies, collagen and native tendon's D-spacings were found to be ~ 61-62 nm (based on $d=2\pi/q$) [157]. Not many studies have seen any two-dimensional SAXS analysis on collagen mimetic peptides to study the D banding periods, this research is the first initiative to study the (GPO) self assemblies. In chapter 4, D-spacing in $(GPO)_{10}$ and $(GPO)_7$ peptides were calculated to be exactly the same, 12.26 Å, indicating similarities in packing for both triple helices. The D-spacing in the $(GOO)_{10}$ peptides was ~12.36 Å, that suggests a different staggered pattern for the helical arrangement and hence would yield a different supramolecular assembly pattern

7.4 Effect of heat on supramolecular assemblies

Heat treatment was seen to cause decreased size of aggregations. This was confirmed by DLS and AFM. AFM analysis on $(GPO)_7$ peptides showed a reduction in assembly heights after heating, when analysed with buffers like acetic acid buffer and high salt PBS conditions.

The heated supramolecular assemblies and collagen peptide forming crystals have recorded peak shifts to higher wavenumbers from the IR spectroscopic experiments in this research. It was speculated that the heat treatments could have resulted in H-bond breakage for all the peptides.

7.4.1 Analysis of supramolecular assemblies by transmission electron microscopy

From studies of Kar et al., $(GPO)_{10}$ peptides show filamentous supramolecular assemblies, which form as the next step of winding back of unfolded helices during cooling and it was predicted that supramolecular assemblies form at a temperature just below the re-naturation temperature [48]. Formation of possible nano-rods for $(GPO)_7$ peptides was seen for the first time in this thesis, when peptides were subjected to heat treatments till 37 °C at a concentration of 7 mg/ml for 8 minutes and were cooled down. 37 °C is the near melting point of helices for $(GPO)_7$ peptides. These also resembled nano-rods formed by peptides lacking hydroxyproline in the work done by Kiick et al. [60].

$(GPO)_3$ peptides were seen to form aggregations, when subjected to heat treatment till 70°C at a concentration of 7 mg/ml for 8 minutes and then cooled down. At the same conditions, $(GPO)_7$ and $(GPO)_{10}$ peptides recorded supramolecular assemblies.

7.4.2 Effect of electrolytes, based on pH and ionic strength

Effects of buffer conditions on the in vitro self-assembly of collagen have been studied before [160] [159]. In previous research, collagen type 1 were assembled in various buffers solutions and structurally investigated by electron microscopy [161-163] [159]. Collagen has also been studied in various buffer conditions with AFM [159, 164, 165]. In this thesis, study of self-assembly of GPO peptides in various buffers has been performed for the first time and high salt PBS environment was found to be most supportive for assembly formations.

7.5 Understanding (GPO) triple helix formation by X-ray crystallography

The crystal structure of the $(GPO)_{10}$ collagen model peptide has been solved at 0.89 Å resolution and atomic details of model triple helices, arranged in a crystalline pseudo-hexagonal network have been understood. Along with confirmation of generic features of model triple helices, such as abundant intra and inter-strand hydrogen bond networks within one triple helix, the localization of glycine residues in the hydrophobic core of the triple stranded structure and the direct hydrogen bonding between triple helices via hydroxyproline side-chains has been established. The other findings are that the proline rings can adopt two stable conformational states, the helical pitch varies around a 7/2 helix and that the stabilization of the crystalline network is water-mediated via hydroxyproline hydration, which bridges the hydrogen bonding of helices.

7.6 Racemic $(GPO)_5$-Gly

Successful racemic crystallization of the smallest possible collagen model peptide of 16 residues, $(GPO)_5$-Gly was achieved. It diffracted at both 1 Å and 1.5 Å resolutions. However, the structure could not be solved.

Melting point of the L-$(GPO)_5$-Gly was determined to be ~35 °C. Isodichroic point was found to be 200 nm. Similar to other GPO peptides, heat-treating these peptides at 70 °C for 8 minutes and cooling down, brought about reduction in size of the aggregates.

7.7 Impact of oxidations on self-assembly

$(GOO)_{10}$ peptides were chosen because presence of hydroxyproline enhances the self-assembly of collagen model peptide chains [60, 201]. Detailed differences in first order assembly, triple helix formation and the higher order assembly, supramolecular assembly formation for $(GPO)_{10}$ and $(GOO)_{10}$ have been examined in chapter 6. Higher melting point and higher stability of the $(GOO)_{10}$ peptides have been established.

7.7.1 Analysis of $(GOO)_{10,}$ for higher stability

From studies of Hayashi et al.. DSC had broader thermal transitions for $(GOO)_{10}$ as compared to $(GPO)_{10}$ [191]. In their study, with $(GOO)_{10}$, no complete conformation changes were observed, when helix adapted a monomeric non-helical state and $(GOO)_{10}$ was more stable than $(GPP)_{10}$ and $(GPO)_{10}$. The possible reasons supporting this higher stability were (1) steric hindrances imposed by the proline ring (2) the inter-chain hydrogen bonds between GlyNH…OC(Xaa) of adjacent chains and (3) hydroxylation of proline residues in Yaa position to 4 (R)-hydroxyproline, such that either hydroxyl group of 4R Hyp forms water bridges between backbone groups or the inductive effect of hydroxyl group influences the ring puckering of pyrrolidine ring and changes the cis/trans ratio of peptide bond [191].

Reason for signal at 225 nm staying positive through melting of $(GOO)_{10}$ as seen by Hayashi et al. was confirmed to be the polyproline II-like conformation of the monomers, following crystal structure of Schumacher et al. [69] [191]. Before the $(GOO)_{10}$ crystal structure was solved in 2005, there were multiple suggestions reasoning its stability and with the structure, most of the reasons were cancelled out. Regarding puckering, it was proposed that the up-puckered conformation of 4R-hydroxyproline in the Xaa position was different to the down-puckered conformation of the Xaa position in $(GPO)_{10}$ peptides, leading to additional stability. But, even after most of the 4 (R) Hyp residues in the Xaa position took up the up-puckered conformation in $(GOO)_{10}$, there was not much difference in the dihedral angels of this structure and thus, this suggestion towards higher $(GOO)_{10}$ stability was discarded. Regarding symmetry, the super helix symmetry of $(GOO)_{10}$ was found to be 7/2, which was similar to other helical symmetries. Hence the only decisive factor, for $(GOO)_{10}$ being more stable than $(GPO)_{10}$ was intra-strand stabilization of polyproline II helix. In this intra-chain water network, the hydroxyl moieties of the Xaa 4(R) Hyp residues hydrogen bind to the water molecule, which is also hydrogen bind to the carbonyl oxygen of the Yaa 4(R)

Hyp residue, located one triplet NH2-terminal to the Xaa 4 (R) Hyp. Thus, the denatured state of Ac-$(GOO)_{10}$-NH_2 possibly adopting a polyproline II like structure was confirmed [69].

There have not been any further solution studies supporting this intra-chain water network way of stabilization until this thesis has evidenced on the intra-strand water stabilization of polyproline II like structure with infrared spectroscopy on $(GOO)_{10}$ crystals and nano-assemblies.

7.7.1.1 IR Spectroscopic analyses of $(GPO)_{10}$ and $(GOO)_{10}$ crystals and nano-assemblies

$(GPO)_{10}$ shows a pattern of H-bonding in assemblies. This pattern is the inter strand H-bonding network and the peptides show random coils after heating which might have occurred because of triple helical melting. In $(GOO)_{10}$, the triple helix was very prominent for unheated supramolecular assemblies, stayed intact after heat treatment. Polyproline II like structural amide I peaks were seen to remain stable even after heat treatments on crystals and higher stability of $(GOO)_{10}$ was confirmed once again with infrared spectroscopy on heated assemblies

7.8 Characterization of oxidative modifications by redox proteomics

Because of cooking, the sites of hydrothermal insult in collagen peptide chain, which impact on the supramolecular assemblies have been recognized. These are mostly the mid-chain sites. The staggered arrangement of the helices is believed to be the primary reason behind such trend of mid-strand zones getting targeted with hydrothermal insult. The changes in nano-scale assemblies occurring after cooking were also visualized by AFM.

7.9 Future work

From literature until today, racemic crystallization of $(GPO)_5$-Gly has been successfully executed in this project work for the first time and the crystals have diffracted at 1.0 Å resolution. Data re-collection and re-solving it with small molecular crystallography could yield a structural solution for the racemic peptide. This will help us to understand the packing of the smallest triple helical structure until date. Racemic crystallography of the even shorter $(GPO)_3$ peptide of 9 residues is another area that can be explored.

There have been many studies related to understanding the proteins modifications, which occur in foods because of processing [202]. There are still some facts unknown in bimolecular protein-protein interactions in re-constituted foods. Redox proteomic profiling

approach has been recently used to understand amino acid modifications by environment and by processing of protein based foods, skin and textiles [8].

To get further insights into the texture of meat derived food products, further studies on oxidative modifications of shorter peptides like $(GPO)_5$ and $(GPO)_3$ can provide further answers. For example (1) whether mid strand sites in a peptide chains are more prone to hydrothermal insult as compared to the end sites, under cooking conditions of various temperatures (2) whether the nature of hydrothermal insults affects the supramolecular assemblies etc.

There have been significant advances in studying primary, secondary, tertiary and quaternary structures of food proteins in recent years but still, there have been fundamental gaps in understanding the interactions of proteins and peptides in whole and reconstituted foods [8, 195]. Controlling the macromolecular interactions between food components has also been a critical consideration in designing functional foods [203]. This thesis paves a way in understanding self-assembling properties of collagen peptides to improve the food matrices, functional foods and meat derived food products.

From literature, evaluation of proteins in food materials after food processing have assisted in understanding how and when protein modifications occur during commercial and domestic food processes [202] [204]. In this thesis, apart from further bridging gaps in the study of oxidative modifications in food industry, this research study on the self assembly of collagen mimetic peptides also holds significance and might find uses in medical applications like drug delivery, tissue engineering and biomedical applications etc. Role of collagen in tissue regeneration is well known and it has an association with numerous pathological conditions. CMPs can pave a way for producing self-assembled constructs similar to tissues, in the field of biomedical engineering. Self-assembly of CMPs will also prove significant in developing collagen-targeting systems in the areas of biomedicine.

8 References

1. Han, R., et al., *Assessment of prokaryotic collagen-like sequences derived from streptococcal Scl1 and Scl2 proteins as a source of recombinant GXY polymers.* Applied Microbiology and Biotechnology, 2006. **72**(1): p. 109-15.
2. Weston, A., R. Rogers, and T. Althen, *Review: the role of collagen in meat tenderness.* Professional Animal Scientist, 2002. **18**(2): p. 107-111.
3. Khoshnoodi, J., et al., *Molecular recognition in the assembly of collagens: terminal noncollagenous domains are key recognition modules in the formation of triple helical protomers.* Journal of Biological Chemistry, 2006. **281**(50): p. 38117-21.
4. Lynn, A., I. Yannas, and W. Bonfield, *Antigenicity and immunogenicity of collagen.* Journal of Biomedical Materials Research Part B: Applied Biomaterials, 2004. **71**(2): p. 343-354.
5. Sakaguchi, M., et al., *IgE reactivity to α1 and α2 chains of bovine type 1 collagen in children with bovine gelatin allergy.* Journal of Allergy and Clinical Immunology, 1999. **104**(3): p. 695-699.
6. Cejas, M.A., et al., *Collagen-related peptides: Self-assembly of short, single strands into a functional biomaterial of micrometer scale.* Journal of the American Chemical Society, 2007. **129**(8): p. 2202-2203.
7. Whipple, G., et al., *Evaluation of attributes that affect longissimus muscle tenderness in Bos taurus and Bos indicus cattle.* Journal of animal science, 1990. **68**(9): p. 2716-2728.
8. Grosvenor, A.J., J.D. Morton, and J.M. Dyer, *Proteomic characterisation of hydrothermal redox damage.* Journal of the Science of Food and Agriculture, 2011.
9. Eyre, D.R., M.A. Paz, and P.M. Gallop, *Cross-linking in collagen and elastin.* Annual Review of Biochemistry, 1984. **53**(1): p. 717-748.
10. Purchas, R., et al., *Concentrations in beef and lamb of taurine, carnosine, coenzyme Q10, and creatine.* Meat Science, 2004. **66**(3): p. 629-637.
11. Mitchell, M., *Carnitine metabolism in human subjects. II. values of carnitine in biological fluids and tissues of" normal" subjects.* American Journal of Clinical Nutrition, 1978. **31**(3): p. 481.
12. Shimada, K., et al., *Species and muscle differences in L-carnitine levels in skeletal muscles based on a new simple assay.* Meat Science, 2004. **68**(3): p. 357-362.
13. Decker, E.A., S.A. Livisay, and S. Zhou, *Mechanism of endogenous skeletal muscle antioxidants: chemical and physical aspects.* Antioxidants in Muscle Foods: Nutritional Strategies to Improve Quality, 2000: p. 25-60.
14. Estévez, M., *Protein carbonyls in meat systems: a review.* Meat Science, 2011.
15. Lund, M.N., et al., *Protein oxidation in muscle foods: a review.* Molecular Nutrition & Food Research, 2011. **55**(1): p. 83-95.
16. Traore, S., et al., *Effect of heat treatment on protein oxidation in pig meat.* Meat Science, 2011.
17. Bailey, A., *The chemistry of collagen cross-links and their role in meat texture.* Proceedings Annual Reciprocal Meat Conference of the American Meat Science Association, 1989. **42**: p. 127-135.
18. Eyre, D.R. and J.J. Wu, *Collagen cross-links.* Collagen, 2005: p. 207-229.
19. Robins, S.P. and A.J. Bailey, *The chemistry of the collagen cross-links. The mechanism of stabilization of the reducible intermediate cross-links.* Biochemical Journal, 1975. **149**(2): p. 381.
20. Cheng, Q. and D.-W. Sun, *Factors affecting the water holding capacity of red meat products: a review of recent research advances.* Critical Reviews in Food Science and Nutrition, 2008. **48**(2): p. 137-159.

21. Lepetit, J., *A theoretical approach of the relationships between collagen content, collagen cross-links and meat tenderness.* Meat Science, 2007. **76**(1): p. 147-159.
22. Bond, J., L. Can, and R. Warner, *The effect of exercise stress, adrenaline injection and electrical stimulation on changes in quality attributes and proteins in Semimembranosus muscle of lamb.* Meat Science, 2004. **68**(3): p. 469-477.
23. Laakkonen, E., *Factors affecting tenderness during heating of meat.* Advances in food research, 1973. **20**: p. 257.
24. Gelse, K., E. Pöschl, and T. Aigner, *Collagens—structure, function, and biosynthesis.* Advanced Drug Delivery Reviews, 2003. **55**(12): p. 1531-1546.
25. Shimokomaki, M., D. Elsden, and A. Bailey, *Meat tenderness: age related changes in bovine intramuscular collagen.* Journal of Food Science, 1972. **37**(6): p. 892-896.
26. Harkness, R.D., *Biological functions of collagen* Biological Reviews, 1961. **36**(4): p. 399-455.
27. McCormick, R., *Extracellular modifications to muscle collagen: implications for meat quality.* Poultry Science, 1999. **78**(5): p. 785-791.
28. Khoshnoodi, J., et al., *Molecular recognition in the assembly of collagens: terminal noncollagenous domains are key recognition modules in the formation of triple helical protomers.* J Biol Chem, 2006. **281**(50): p. 38117-21.
29. Light, N. and A.E. Champion, *Characterization of muscle epimysium, perimysium and endomysium collagens.* Biochemical Journal, 1984. **219**(3): p. 1017.
30. Alberts, B., et al., *Molecular Biology of the Cell (3rd edn).* Trends in Biochemical Sciences, 1995. **20**(5): p. 210-210.
31. Kadler, K.E., et al., *Collagen fibril formation.* Biochemical Journal, 1996. **316**(Pt 1): p. 1.
32. Fallas, J.A., L.E.R. O'Leary, and J.D. Hartgerink, *Synthetic collagen mimics: self-assembly of homotrimers, heterotrimers and higher order structures.* Chemical Society Reviews, 2010. **39**(9): p. 3510-3527.
33. Shoulders, M.D. and R.T. Raines, *Collagen structure and stability.* Annual Review of Biochemistry, 2009. **78**: p. 929.
34. Buehler, M.J., *Nature designs tough collagen: explaining the nanostructure of collagen fibrils.* Proceedings of the National Academy of Sciences, USA, 2006. **103**(33): p. 12285-12290.
35. Miyahara, M., et al., *Formation of collagen fibrils by enzymic cleavage of precursors of type I collagen in vitro.* Journal of Biological Chemistry, 1984. **259**(15): p. 9891-9898.
36. Miyahara, M., F. Njieha, and D. Prockop, *Formation of collagen fibrils in vitro by cleavage of procollagen with procollagen proteinases.* Journal of Biological Chemistry, 1982. **257**(14): p. 8442-8448.
37. Kadler, K.E., Y. Hojima, and D. Prockop, *Assembly of collagen fibrils de novo by cleavage of the type I pC-collagen with procollagen C-proteinase. Assay of critical concentration demonstrates that collagen self-assembly is a classical example of an entropy-driven process.* Journal of Biological Chemistry, 1987. **262**(32): p. 15696.
38. Vogel, B., et al., *A substitution of cysteine for glycine 748 of the alpha 1 chain produces a kink at this site in the procollagen I molecule and an altered N-proteinase cleavage site over 225 nm away.* Journal of Biological Chemistry, 1988. **263**(35): p. 19249.
39. Bhattacharjee, A. and M. Bansal, *Collagen structure: the madras triple helix and the current scenario.* Lubmb Life, 2005. **57**(3): p. 161-172.
40. Klug, W.S. and R. Michael, *Cummings. Concepts of Genetics*. 1997: Prentice Hall.
41. Eyre, D.R. and J.-J. Wu, *Collagen Cross-Links*, in *Collagen: Primer in Structure, Processing and Assembly*, J. Brinckmann, H. Notbohm, and P.K. Müller, Editors. 2005, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 207-229.
42. Eyre, D.R. and M.J. Glimcher, *Reducible crosslinks in hydroxylysine-deficient collagens of a heritable disorder of connective tissue.* Proceedings of the National Academy of Sciences, USA, 1972. **69**(9): p. 2594-2598.

43. Rele, S., et al., *D-periodic collagen-mimetic microfibers.* Journal of the American Chemical Society, 2007. **129**(47): p. 14780-14787.
44. Castelletto, V., et al., *Fibrillar superstructure from extended nanotapes formed by a collagen-stimulating peptide.* Chemical Communications, 2010. **46**(48): p. 9185-9187.
45. Kotch, F.W. and R.T. Raines, *Self-assembly of synthetic collagen triple helices.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(9): p. 3028-3033.
46. Li, M.H., et al., *Two-dimensional NMR assignments and conformation of (Pro-Hyp-Gly) 10 and a designed collagen triple-helical peptide.* Biochemistry, 1993. **32**(29): p. 7377-7387.
47. Kramer, R.Z., et al., *X-ray crystallographic determination of a collagen-like peptide with the repeating sequence (Pro-Pro-Gly).* Journal of Molecular Biology, 1998. **280**(4): p. 623-638.
48. Kar, K., et al., *Self-association of collagen triple helic peptides into higher order structures.* Journal of Biological Chemistry, 2006. **281**(44): p. 33283-33290.
49. Kar, K., et al., *Self-association of collagen triple helic peptides into higher order structures.* J Biol Chem, 2006. **281**(44): p. 33283-90.
50. Woolfson, D.N., *The Design of coiled-coil structures and assemblies*, in *Advances in Protein Chemistry*, A.D.P. David and M.S. John, Editors. 2005, Academic Press. p. 79-112.
51. Suto, K. and H. Noda, *Conformational change of the triple-helical structure. IV. Kinetics of the helix-folding of (Pro-Pro-Gly)n (n equals 10, 12, and 15).* Biopolymers, 1974. **13**(12): p. 2477-88.
52. Nichol, J.W. and A. Khademhosseini, *Modular tissue engineering: engineering biological tissues from the bottom up.* Soft Matter, 2009. **5**(7): p. 1312-1319.
53. O'Leary, L.E.R., et al., *Multi-hierarchical self-assembly of a collagen mimetic peptide from triple helix to nanofibre and hydrogel.* Nature Chemistry, 2011. **3**(10): p. 821-828.
54. Kuznetsova, N., S. Chi, and S. Leikin, *Sugars and polyols inhibit fibrillogenesis of type I collagen by disrupting hydrogen-bonded water bridges between the helices.* Biochemistry, 1998. **37**(34): p. 11888-11895.
55. Cejas, M.A., et al., *Thrombogenic collagen-mimetic peptides: Self-assembly of triple helix-based fibrils driven by hydrophobic interactions.* Proceedings of the National Academy of Sciences, 2008. **105**(25): p. 8513.
56. Cejas, M.A., et al., *Thrombogenic collagen-mimetic peptides: self-assembly of triple helix-based fibrils driven by hydrophobic interactions.* Proceedings of the National Academy of Sciences, USA, 2008. **105**(25): p. 8513-8518.
57. Rele, S., et al., *D-periodic collagen-mimetic microfibers.* Journal of the American Chemical Society, 2007. **129**(47): p. 14780-14787.
58. Zhu, Y., et al., *Type IIA procollagen containing the cysteine-rich amino propeptide is deposited in the extracellular matrix of prechondrogenic tissue and binds to TGF-β1 and BMP-2.* Journal of Cell Biology, 1999. **144**(5): p. 1069-1080.
59. Shacter, E., *Qualtification and significance of protein oxidation in biological samples.* Drug Metabolism Reviews, 2000. **32**(3-4): p. 307-326.
60. Krishna, O.D. and K.L. Kiick, *Supramolecular assembly of electrostatically stabilized, hydroxyproline-lacking collagen-mimetic peptides.* Biomacromolecules, 2009. **10**(9): p. 2626-2631.
61. Vitagliano, L., et al., *Structural bases of collagen stabilization induced by proline hydroxylation.* Biopolymers, 2001. **58**(5): p. 459-464.
62. Kramer, R.Z., et al., *The crystal and molecular structure of a collagen-like peptide with a biologically relevant sequence.* Journal of Molecular Biology, 2001. **311**(1): p. 131-147.
63. Hodges, J.A. and R.T. Raines, *Stereoelectronic effects on collagen stability: the dichotomy of 4-fluoroproline diastereomers.* Journal of the American Chemical Society, 2003. **125**(31): p. 9262-9263.

64. Gauba, V. and J.D. Hartgerink, *Self-assembled heterotrimeric collagen triple helices directed through electrostatic interactions.* Journal of the American Chemical Society, 2007. **129**(9): p. 2683-2690.
65. Russell, L.E., J.A. Fallas, and J.D. Hartgerink, *Selective assembly of a high stability AAB collagen heterotrimer.* Journal of the American Chemical Society, 2010. **132**(10): p. 3242-3243.
66. Kawahara, K., et al., *Effect of hydration on the stability of the collagen-like triple-helical structure of [4 (R)-hydroxyprolyl-4 (R)-hydroxyprolylglycine] 10.* Biochemistry, 2005. **44**(48): p. 15812-15822.
67. Berisio, R., et al., *Imino acids and collagen triple helix stability: Characterization of collagen-like polypeptides containing Hyp-Hyp-Gly sequence repeats.* Journal of the American Chemical Society, 2004. **126**(37): p. 11402-11403.
68. Emsley, J., et al., *Structural basis of collagen recognition by integrin α2β1.* Cell, 2000. **101**(1): p. 47-56.
69. Schumacher, M., K. Mizuno, and H.P. Bachinger, *The crystal structure of the collagen-like polypeptide (Glycyl-4(R)-hydroxyprolyl-4(R)-hydroxyprolyl)(9) at 1.55 angstrom resolution shows up-puckering of the proline ring in the Xaa position.* Journal of Biological Chemistry, 2005. **280**(21): p. 20397-20403.
70. Okuyama, K., et al., *Crystal and molecular structure of a collagen-like polypeptide (Pro-Pro-Gly)$_{10}$.* Journal of Molecular Biology, 1981. **152**(2): p. 427-443.
71. Fraser, R., et al., *Molecular conformation and packing in collagen fibrils.* Journal of Molecular Biology, 1983. **167**(2): p. 497-521.
72. Roy, S., et al., *Structure prediction and functional characterization of secondary metabolite proteins of Ocimum.* Bioinformation, 2011. **6**(8): p. 315.
73. Rebecchi, K.R., et al., *A general protease digestion procedure for optimal protein sequence coverage and PTM analysis of recombinant glycoproteins: Application to the characterization of hLOXL2 glycosylation.* Analytical Chemistry, 2011. **83**(22): p. 8484-8491.
74. Chung, L., et al., *Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis.* The European Molecular Biology Organization Journal, 2004. **23**(15): p. 3020-3030.
75. Silva, T.H., et al., *Marine origin collagens and its potential applications.* Marine Drugs, 2014. **12**(12): p. 5881-5901.
76. Gobom, J., et al., *α-Cyano-4-hydroxycinnamic acid affinity sample preparation. A protocol for MALDI-MS peptide analysis in proteomics.* Analytical Chemistry, 2001. **73**(3): p. 434-438.
77. Zhukov, A., et al., *Integration of surface plasmon resonance with mass spectrometry: automated ligand fishing and sample preparation for MALDI MS using a biacore 3000 biosensor.* Journal of Biomolecular Techniques 2004. **15**(2): p. 112-119.
78. Dyer, J.M., et al., *Proteomic evaluation and location of UVB-induced photo-oxidation in wool.* Journal of Photochemistry and Photobiology B: Biology, 2010. **98**(2): p. 118-27.
79. Engel, J. and H. Bächinger, *Structure, Stability and Folding of the Collagen Triple Helix*, in *Collagen*, J. Brinckmann, H. Notbohm, and P.K. Müller, Editors. 2005, Springer Berlin Heidelberg. p. 7-33.
80. Pelc, D., et al., *Role of microscopic phase separation in gelation of aqueous gelatin solutions.* Soft Matter, 2014. **10**(2): p. 348-356.
81. Brodsky, B., et al., *Triple-helical peptides: an approach to collagen conformation, stability, and self-association.* Biopolymers, 2008. **89**(5): p. 345-353.
82. Kar, K., et al., *Aromatic interactions promote self-association of collagen triple-helical peptides to higher-order structures.* Biochemistry, 2009. **48**(33): p. 7959-7968.
83. O'Neill, M.J., *The analysis of a temperature-controlled scanning calorimeter.* Analytical Chemistry, 1964. **36**(7): p. 1238-1245.

84. Chiu, M.H. and E.J. Prenner, *Differential scanning calorimetry: An invaluable tool for a detailed thermodynamic characterization of macromolecules and their interactions.* Journal of Pharmacy and Bioallied Sciences, 2011. **3**(1): p. 39-59.
85. Wille, H., et al., *Alzheimer-like paired helical filaments and antiparallel dimers formed from microtubule-associated protein tau In vitro.* Journal of Cell Biology, 1992. **118**(3): p. 573-584.
86. Krimm, S. and J. Bandekar, *Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins*, in *Advances in Protein Chemistry*, J.T.E. C.B. Anfinsen and M.R. Frederic, Editors. 1986, Academic Press. p. 181-364.
87. Doyle, B.B., E. Bendit, and E.R. Blout, *Infrared spectroscopy of collagen and collagen like polypeptides.* Biopolymers, 1975. **14**(5): p. 937-957.
88. van Grondelle, W., et al., *Lamination and spherulite-like compaction of a hormone's native amyloid-like nanofibrils: spectroscopic insights into key interactions.* Faraday Discussions, 2013. **166**(0): p. 163-180.
89. Belbachir, K., et al., *Collagen types analysis and differentiation by FTIR spectroscopy.* Analytical and Bioanalytical Chemistry, 2009. **395**(3): p. 829-837.
90. Chernoff, E.A. and D.A. Chernoff, *Atomic force microscope images of collagen fibers.* Journal of Vacuum Science & Technology A, 1992. **10**(4): p. 596-599.
91. Gill, S.C. and P.H. Von Hippel, *Calculation of protein extinction coefficients from amino acid sequence data.* Analytical Biochemistry, 1989. **182**(2): p. 319-326.
92. Gill, S.C. and P.H. von Hippel, *Calculation of protein extinction coefficients from amino acid sequence data.* Analytical Biochemistry, 1989. **182**(2): p. 319-26.
93. Lukaski, H.C., *Methods for the assessment of human body composition: traditional and new.* The American Journal of Clinical Nutrition, 1987. **46**(4): p. 537-56.
94. Engel, J. and H.P. Bächinger, *Structure, stability and folding of the collagen triple helix*
collagen, in *Topics in Current Chemistry*, J. Brinckmann, H. Notbohm, and P.K. Müller, Editors. 2005, Springer Berlin / Heidelberg. p. 7-33.
95. Bergfors, T., *Seeds to crystals.* Journal of Structural Biology, 2003. **142**(1): p. 66-76.
96. Matthews, B.W., *Solvent content of protein crystals.* Journal of Molecular Biology, 1968. **33**(2): p. 491-497.
97. Ringe, D. and G.A. Petsko, *Study of protein dynamics by X-ray diffraction.* Methods in Enzymology, 1986. **131**: p. 389-433.
98. Kabsch, W., *XDS.* Acta Crystallographica Section D: Biological Crystallography, 2010. **66**: p. 125-132.
99. Sheldrick, G.M., *Experimental phasing with SHELXC/D/E: combining chain tracing with density modification.* Acta Crystallographica Section D: Biological Crystallography, 2010. **66**: p. 479-485.
100. Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method.* Acta Crystallographica Section D: Biological Crystallography, 1997. **53**: p. 240-255.
101. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallographica Section D: Biological Crystallography, 2004. **60**: p. 2126–2132.
102. Kabsch, W. and C. Sander, *Dictionary of protein secondary structure: pattern recognition of hydrogen bonded and geometrical features.* Biopolymers, 1983. **22**: p. 2577-2637.
103. Collaborative, C.P., *The CCP4 suite: programs for protein crystallography.* Acta crystallographica. Section D, Biological crystallography, 1994. **50**(Pt 5): p. 760.
104. Zhang, S., et al., *Design of nanostructured biological materials through self-assembly of peptides and proteins.* Current Opinion in Chemical Biology, 2002. **6**(6): p. 865-871.
105. Paradís, M., et al., *RADA-16: A Tough Peptide–Strategies for Synthesis and Purification.* European Journal of Organic Chemistry, 2013. **2013**(26): p. 5871-5878.
106. Cen, L., et al., *Collagen tissue engineering: development of novel biomaterials and applications.* Pediatric Research, 2008. **63**(5): p. 492-496.

107. Place, E.S., N.D. Evans, and M.M. Stevens, *Complexity in biomaterials for tissue engineering.* Nature Materials, 2009. **8**(6): p. 457-470.
108. Brodsky, B., et al., *Triple-helical peptides: An approach to collagen conformation, stability, and self-association.* Biopolymers, 2008. **89**(5): p. 345-353.
109. Persikov, A.V., Y. Xu, and B. Brodsky, *Equilibrium thermal transitions of collagen model peptides.* Protein Science, 2004. **13**(4): p. 893-902.
110. Kishimoto, T., et al., *Synthesis of poly(Pro–Hyp–Gly)n by direct polycondensation of (Pro–Hyp–Gly)n, where n = 1, 5, and 10, and stability of the triple-helical structure.* Biopolymers, 2005. **79**(3): p. 163-172.
111. Martin, R., L. Waldmann, and D.L. Kaplan, *Supramolecular assembly of collagen triblock peptides.* Biopolymers, 2003. **70**(4): p. 435-44.
112. Inouye, K., et al., *Synthesis and physical properties of (hydroxyproline-proline-glycine)10: Hydroxyproline in the X-position decreases the melting temperature of the collagen triple helix.* Archives of Biochemistry and Biophysics, 1982. **219**(1): p. 198-203.
113. Holmgren, S.K., et al., *A hyperstable collagen mimic.* Chemistry & Biology, 1999. **6**(2): p. 63-70.
114. Okuyama, K., et al., *Crystal structures of collagen model peptides with Pro-Hyp-Gly repeating sequence at 1.26 Å resolution: Implications for proline ring puckering.* Peptide Science, 2004. **76**(5): p. 367-377.
115. Liu, X., et al., *Direct NMR measurement of the folding kinetics of a trimeric peptide.* Biochemistry, 1996. **35**(14): p. 4306-4313.
116. Greenfield, N.J., *Using circular dichroism spectra to estimate protein secondary structure.* Nature Protocols, 2006. **1**(6): p. 2876-2890.
117. Besley, N.A. and J.D. Hirst, *Theoretical studies toward quantitative protein circular dichroism calculations.* Journal of the American Chemical Society, 1999. **121**(41): p. 9636-9644.
118. Janes, R.W., et al., *An introduction to circular dichroism and synchrotron radiation circular dichroism spectroscopy*. Modern techniques for circular dichroism and synchrotron radiation circular dichroism spectroscopy. IOS Press, Amsterdam. 2009. 1-18.
119. Venyaminov, S.Y. and J.T. Yang, *Determination of protein secondary structure*. Circular dichroism and the conformational analysis of biomolecules. 1996: Springer. 69-107.
120. Persikov, A.V., J.A. Ramshaw, and B. Brodsky, *Prediction of collagen stability from amino acid sequence.* Journal of Biological Chemistry, 2005. **280**(19): p. 19343-19349.
121. Yang, W., et al., *Gly-Pro-Arg Confers Stability Similar to Gly-Pro-Hyp in the Collagen Triple-helix of Host-Guest Peptides.* Journal of Biological Chemistry, 1997. **272**(46): p. 28837-28840.
122. Engel, J. and H.P. Bächinger, *Structure, stability and folding of the collagen triple helix*, in *Collagen*. 2005, Springer. p. 7-33.
123. Buhot, A. and A. Halperin, *Extension behavior of helicogenic polypeptides.* Macromolecules, 2002. **35**(8): p. 3238-3252.
124. Streicher, W.W. and G.I. Makhatadze, *Unfolding thermodynamics of Trp-cage, a 20 residue miniprotein, studied by differential scanning calorimetry and circular dichroism spectroscopy.* Biochemistry, 2007. **46**(10): p. 2876-2880.
125. Yu, L., et al., *Thermochemistry and conformational polymorphism of a hexamorphic crystal system.* Journal of the American Chemical Society, 2000. **122**(4): p. 585-591.
126. Bruylants, G., J. Wouters, and C. Michaux, *Differential scanning calorimetry in life science: thermodynamics, stability, molecular recognition and application in drug design.* Current Medicinal Chemistry, 2005. **12**(17): p. 2011-2020.
127. Privalov, P.L. and S.A. Potekhin, *[2] Scanning microcalorimetry in studying temperature-induced changes in proteins.* Methods in Enzymology, 1986. **131**: p. 4-51.
128. Biltonen, R.L., E. Freire, and J.F. Brandts, *Thermodynamic characterization of conformational states of biological macromolecules using differential scanning calorimetry.* CRC Critical Reviews in Biochemistry, 1978. **5**(2): p. 85-124.

129. Kurzynski, M., *Chemical Reactions.* The Thermodynamic Machinery of Life, 2006: p. 141-171.
130. Pasini, D. and A. Kraft, *Supramolecular self-assembly of fibres.* Current Opinion in Solid State and Materials Science, 2004. **8**(2): p. 157-163.
131. Collier, J.H. and P.B. Messersmith, *Self-Assembling Polymer–Peptide Conjugates: Nanostructural Tailoring.* Advanced Materials, 2004. **16**(11): p. 907-910.
132. Shibata-Seki, T., et al., *Application of atomic force microscopy to protein anatomy:. Imaging of supramolecular structures of self-assemblies formed from synthetic peptides.* Applied Physics A: Materials Science & Processing, 1998. **66**: p. S625-S629.
133. Kar, K., Y.H. Wang, and B. Brodsky, *Sequence dependence of kinetics and morphology of collagen model peptide self-assembly into higher order structures.* Protein Science, 2008. **17**(6): p. 1086-1095.
134. Schneider, J.P., et al., *Responsive hydrogels from the intramolecular folding and self-assembly of a designed peptide.* Journal of the American Chemical Society, 2002. **124**(50): p. 15030-15037.
135. Shibata-Seki, T., et al., *Application of atomic force microscopy to protein anatomy:. Imaging of supramolecular structures of self-assemblies formed from synthetic peptides.* Applied Physics A: Materials Science & Processing, 1998. **66**(7): p. S625.
136. Mammadov, R., et al., *Microscopic characterization of peptide nanostructures.* Micron, 2012. **43**(2-3): p. 69-84.
137. Chamberlain, A.K., et al., *Ultrastructural Organization of Amyloid Fibrils byAtomic Force Microscopy.* Biophysical Journal, 2000. **79**(6): p. 3282-3293.
138. Kar, K., Y.-H. Wang, and B. Brodsky, *Sequence dependence of kinetics and morphology of collagen model peptide self-assembly into higher order structures.* Protein Science : A Publication of the Protein Society, 2008. **17**(6): p. 1086-1095.
139. Leikin, S., D. Rau, and V. Parsegian, *Temperature-favoured assembly of collagen is driven by hydrophilic not hydrophobic interactions.* Nature Structural Biology, 1995. **2**(3): p. 205-210.
140. Przybyla, D.E. and J. Chmielewski, *Higher-Order Assembly of Collagen Peptides into Nano- and Microscale Materials.* Biochemistry, 2010. **49**(21): p. 4411-4419.
141. Leikina, E., et al., *Type I collagen is thermally unstable at body temperature.* Proceedings of the National Academy of Sciences, 2002. **99**(3): p. 1314-1318.
142. Bryan, M.A., et al., *FTIR studies of collagen model peptides: complementary experimental and simulation approaches to conformation and unfolding.* Journal of the American Chemical Society, 2007. **129**(25): p. 7877-7884.
143. Smejkal, B., et al., *Protein crystallization in stirred systems—scale-up via the maximum local energy dissipation.* Biotechnology and Bioengineering, 2013. **110**(7): p. 1956-1963.
144. Prystupa, D. and A. Donald, *Infrared study of gelatin conformations in the gel and sol states.* Polymer Gels and Networks, 1996. **4**(2): p. 87-110.
145. Doyle, B.B., E. Bendit, and E.R. Blout, *Infrared spectroscopy of collagen and collagen-like polypeptides.* Biopolymers, 1975. **14**(5): p. 937-957.
146. Lazarev, Y.A., B. Grishkovsky, and T. Khromova, *Amide I band of IR spectrum and structure of collagen and related polypeptides.* Biopolymers, 1985. **24**(8): p. 1449-1478.
147. de Campos Vidal, B. and M.L.S. Mello, *Collagen type I amide I band infrared spectroscopy.* Micron, 2011. **42**(3): p. 283-289.
148. Muyonga, J.H., C.G.B. Cole, and K.G. Duodu, *Fourier transform infrared (FTIR) spectroscopic study of acid soluble collagen and gelatin from skins and bones of young and adult Nile perch (Lates niloticus).* Food Chemistry, 2004. **86**(3): p. 325-332.
149. Camacho, N.P., et al., *FTIR microscopic imaging of collagen and proteoglycan in bovine cartilage.* Biopolymers, 2001. **62**(1): p. 1-8.
150. Hu, X., et al., *The Influence of Elasticity and Surface Roughness on Myogenic and Osteogenic-Differentiation of Cells on Silk-Elastin Biomaterials.* Biomaterials, 2011. **32**(34): p. 8979-8989.

151. Jastrzebska, M., et al., *Tannic acid-stabilized pericardium tissue: IR spectroscopy, atomic force microscopy, and dielectric spectroscopy investigations.* Journal of Biomedical Materials Research Part A, 2006. **78A**(1): p. 148-156.
152. Tiong, W.H., et al., *Enhancing amine terminals in an amine-deprived collagen matrix.* Langmuir, 2008. **24**(20): p. 11752-11761.
153. Jackson, M. and H.H. Mantsch, *The use and misuse of FTIR spectroscopy in the determination of protein structure.* Critical Reviews in Biochemistry and Molecular biology, 1995. **30**(2): p. 95-120.
154. Naumann, D., et al., *FT-IR spectroscopy and FT-Raman spectroscopy are powerful analytical tools for the non-invasive characterization of intact microbial cells.* Journal of Molecular Structure, 1995. **347**: p. 399-405.
155. Andrushchenko, V.V., H.J. Vogel, and E.J. Prenner, *Solvent-dependent structure of two tryptophan-rich antimicrobial peptides and their analogs studied by FTIR and CD spectroscopy.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 2006. **1758**(10): p. 1596-1608.
156. Goormaghtigh, E. and V.R. Cabiaux, Jean-Marie, *Secondary structure and dosage of soluble and membrane proteins by attenuated total reflection Fourier-transform infrared spectroscopy on hydrated films.* European Journal of Biochemistry, 1990. **193**(2): p. 409-420.
157. Cheng, X., et al., *An electrochemical fabrication process for the assembly of anisotropically oriented collagen bundles.* Biomaterials, 2008. **29**(22): p. 3278-3288.
158. Jiang, T., et al., *Structurally defined nanoscale sheets from self-assembly of collagen-mimetic peptides.* Journal of the American Chemical Society, 2014. **136**(11): p. 4300-4308.
159. Jiang, F., et al., *Assembly of collagen into microribbons: effects of pH and electrolytes.* Journal of Structural Biology, 2004. **148**(3): p. 268-278.
160. Hattori, S., et al., *Alkali-treated collagen retained the triple helical conformation and the ligand activity for the cell adhesion via α2β1 integrin.* Journal of Biochemistry, 1999. **125**(4): p. 676-684.
161. Kitagawa, M., et al., *An F-box protein, FWD1, mediates ubiquitin-dependent proteolysis of β-catenin.* The EMBO Journal, 1999. **18**(9): p. 2401-2410.
162. Perret, S., et al., *Unhydroxylated triple helical collagen I produced in transgenic plants provides new clues on the role of hydroxyproline in collagen folding and fibril formation.* Journal of Biological Chemistry, 2001. **276**(47): p. 43693-43698.
163. Suzuki, Y., et al., *Interaction of collagen molecules from the aspect of fibril formation: acid-soluble, alkali-treated, and MMP1-digested fragments of type I collagen.* Journal of Biochemistry, 1999. **126**(1): p. 54-67.
164. Agarwal, G., et al., *Binding of discoidin domain receptor 2 to collagen I: an atomic force microscopy investigation.* Biochemistry, 2002. **41**(37): p. 11091-11098.
165. Gale, M., et al., *Sequential assembly of collagen revealed by atomic force microscopy.* Biophysical Journal, 1995. **68**(5): p. 2124.
166. Wilson, D.L., et al., *Surface organization and nanopatterning of collagen by dip-pen nanolithography.* Proceedings of the National Academy of Sciences, 2001. **98**(24): p. 13660-13664.
167. Schön, P.M., M. Gosa, and G.J. Vancso, *Imaging of Mechanical Properties of Soft Matter. From Heterogeneous Polymer Surfaces to Single Biomolecules.* Imaging & microscopy, 2013. **15**: p. 43-46.
168. Landoulsi, J. and V. Dupres, *Direct AFM force mapping of surface nanoscale organization and protein adsorption on an aluminum substrate.* Physical Chemistry Chemical Physics, 2013. **15**(21): p. 8429-8440.
169. Bella, J., et al., *Crystal and molecular structure of a collagen-like peptide at 1.9 A resolution.* Science, 1994. **266**(5182): p. 75-81.
170. Ebert, G., *3 Polypeptide and Proteine.* Biopolymere, 1992.

171. Berisio, R., et al., *Crystal structure of a collagen-like polypeptide with repeating sequence Pro–Hyp–Gly at 1.4 Å resolution: Implications for collagen hydration.* Biopolymers, 2000. **56**(1): p. 8-13.
172. Okuyama, K., et al., *Crystal structures of collagen model peptides with Pro-Hyp-Gly repeating sequence at 1.26 angstrom resolution: Implications for proline ring puckering.* Biopolymers, 2004. **76**(5): p. 367-377.
173. Okuyama, K., et al., *Crystal structure of (Gly-Pro-Hyp)9: Implications for the collagen molecular model.* Biopolymers, 2012. **97**(8): p. 607-616.
174. Okuyama, K., et al., *Helical twists of collagen model peptides.* Peptide Science, 2006. **84**(4): p. 421-432.
175. Smatanová, I.K., *Crystallization of biological macromolecules.* Materials Structure, 2002. **9**(1).
176. Bergfors, T.M., *Protein Crystallization*. 2009: International University Line.
177. Blundell, T.L. and L.N. Johnson, *Protein Crystallography.* 1976.
178. Caffrey, M. and V. Cherezov, *Crystallizing membrane proteins using lipidic mesophases.* Nature Protocols, 2009. **4**(5): p. 706-731.
179. Fallas, J.A. and J.D. Hartgerink, *Computational design of self-assembling register-specific collagen heterotrimers.* Nature Communications, 2012. **3**: p. 1087.
180. Okuyama, K., et al., *High-resolution structures of collagen-like peptides [(Pro-Pro-Gly) 4-Xaa-Yaa-Gly-(Pro-Pro-Gly) 4]: Implications for triple-helix hydration and Hyp (X) puckering.* Biopolymers, 2009. **91**(5): p. 361-372.
181. Kang, Y.K. and H. Young Choi, *Cis–trans isomerization and puckering of proline residue.* Biophysical Chemistry, 2004. **111**(2): p. 135-142.
182. Okuyama, K., et al., *High-Resolution Structures of Collagen-Like Peptides (Pro-Pro-Gly)(4)-Xaa-Yaa-Gly-(Pro-Pro-Gly)(4) : Implications for Triple-Helix Hydration and Hyp(X) Puckering.* Biopolymers, 2009. **91**(5): p. 361-372.
183. Sweeney, S.M., et al., *Defining the domains of type I collagen involved in heparin-binding and endothelial tube formation.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(13): p. 7275-7280.
184. Emsley, J., et al., *Structural basis of collagen recognition by integrin alpha 2 beta 1.* Cell, 2000. **101**(1): p. 47-56.
185. Di Lullo, G.A., et al., *Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen.* Journal of Biological Chemistry, 2002. **277**(6): p. 4223-4231.
186. Okuyama, K., et al., *Revision of collagen molecular structure.* Biopolymers, 2006. **84**(2): p. 181-191.
187. Kramer, R.Z., et al., *Sequence dependent conformational variations of collagen triple-helical structure.* Nature Structural Biology, 1999. **6**(5): p. 454.
188. Boudko, S.P., J. Engel, and H.P. Bächinger, *The crucial role of trimerization domains in collagen folding.* The International Journal of Biochemistry & Cell Biology, 2012. **44**(1): p. 21-32.
189. Kramer, R.Z., et al., *Staggered molecular packing in crystals of a collagen-like peptide with a single charged pair.* Journal of Molecular Biology, 2000. **301**(5): p. 1191-1205.
190. Melacini, G., et al., *Hydration dynamics of the collagen triple helix by NMR.* Journal of Molecular Biology, 2000. **300**(5): p. 1041-1048.
191. Mizuno, K., et al., *Hydroxylation-induced Stabilization of the Collagen triple helix: acetyl-(Glycyl-4(R)-hydroxyprolyl-4(R)-hydroxyprolyl)10-NH2 forms a highly stable triple helix.* Journal of Biological Chemistry, 2004. **279**(36): p. 38072-38078.
192. Inouye, K., et al., *Synthesis and physical properties of (hydroxyproline-proline-glycine) 10: hydroxyproline in the X-position decreases the melting temperature of the collagen triple helix.* Archives of Biochemistry and Biophysics, 1982. **219**(1): p. 198-203.

193. Brodsky, B. and J.A. Ramshaw, *The collagen triple-helix structure.* Matrix Biology, 1997. **15**(8): p. 545-554.
194. Wishart, D.S., *Metabolomics: applications to food science and nutrition research.* Trends in Food Science & Technology, 2008. **19**(9): p. 482-493.
195. Grosvenor, A.J., J.D. Morton, and J.M. Dyer, *Isobaric labeling approach to the tracking and relative quantitation of peptide damage at the primary structural level.* Journal of Agricultural and Food Chemistry, 2010. **58**(24): p. 12672-12677.
196. Frank, S., et al., *Stabilization of short collagen-like triple helices by protein engineering1.* Journal of Molecular Biology, 2001. **308**(5): p. 1081-1089.
197. Brahmachari, S.K., et al., *Structural investigations on poly (4-hydroxy-L-proline). 2. Physicochemical studies.* Macromolecules, 1979. **12**(1): p. 23-28.
198. Gill, P., T.T. Moghadam, and B. Ranjbar, *Differential Scanning Calorimetry Techniques: Applications in Biology and Nanoscience.* Journal of Biomolecular Techniques : JBT, 2010. **21**(4): p. 167-193.
199. Adochitei, A. and G. Drochioiu, *Rapid characterization of peptide secondary structure by FT-IR spectroscopy.* Romanian Journal of Chemistry, 2011. **56**(8): p. 783-791.
200. Gurry, T., P.S. Nerenberg, and C.M. Stultz, *The contribution of interchain salt bridges to triple-helical stability in collagen.* Biophysical journal, 2010. **98**(11): p. 2634-2643.
201. Krishna, O.D. and K.L. Kiick, *Protein- and Peptide-modified synthetic polymeric biomaterials.* Biopolymers, 2010. **94**(1): p. 32-48.
202. Friedman, M., *Chemistry, biochemistry, and safety of acrylamide. A review.* Journal of Agricultural and Food Chemistry, 2003. **51**(16): p. 4504-4526.
203. Asakura, S. and F. Oosawa, *On interaction between two bodies immersed in a solution of macromolecules.* Chemical Physics, 1954(22): p. 1255-1256.
204. Gallier, S., et al., *In vivo digestion of bovine milk fat globules: Effect of processing and interfacial structural changes. II. Upper digestive tract digestion.* Food Chemistry, 2013. **141**(3): p. 3215-3223.